National Library
of Canada

Acquisitions and
Bibliographic Services Branch

395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale
du Canada

Direction des acquisitions et
des services bibliographiques

395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Canada

UNIVERSITY OF ALBERTA

BODY COMPOSITION AND SELF-ESTEEM: A VALIDATION OF SELECTED SELF-ESTEEM MEASURES

BY

RICHARD BALLARD FLETCHER

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfilment of the requirements for the degree of

MASTER OF SCIENCE

FACULTY OF PHYSICAL EDUCATION AND SPORTS STUDIES

EDMONTON ALBERTA

FALL, 1995

National Library of Canada
Acquisitions and Bibliographic Services Branch
395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada
Direction des acquisitions et des services bibliographiques
395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

THE AUTHOR HAS GRANTED AN IRREVOCABLE NON-EXCLUSIVE LICENCE ALLOWING THE NATIONAL LIBRARY OF CANADA TO REPRODUCE, LOAN, DISTRIBUTE OR SELL COPIES OF HIS/HER THESIS BY ANY MEANS AND IN ANY FORM OR FORMAT, MAKING THIS THESIS AVAILABLE TO INTERESTED PERSONS.

L'AUTEUR A ACCORDE UNE LICENCE IRREVOCABLE ET NON EXCLUSIVE PERMETTANT A LA BIBLIOTHEQUE NATIONALE DU CANADA DE REPRODUIRE, PRETER, DISTRIBUER OU VENDRE DES COPIES DE SA THESE DE QUELQUE MANIERE ET SOUS QUELQUE FORME QUE CE SOIT POUR METTRE DES EXEMPLAIRES DE CETTE THESE A LA DISPOSITION DES PERSONNE INTERESSEES.

THE AUTHOR RETAINS OWNERSHIP OF THE COPYRIGHT IN HIS/HER THESIS. NEITHER THE THESIS NOR SUBSTANTIAL EXTRACTS FROM IT MAY BE PRINTED OR OTHERWISE REPRODUCED WITHOUT HIS/HER PERMISSION.

L'AUTEUR CONSERVE LA PROPRIETE DU DROIT D'AUTEUR QUI PROTEGE SA THESE. NI LA THESE NI DES EXTRAITS SUBSTANTIELS DE CELLE-CI NE DOIVENT ETRE IMPRIMES OU AUTREMENT REPRODUITS SANS SON AUTORISATION.

ISBN 0-612-06470-0

Canada

UNIVERSITY OF ALBERTA

LIBRARY RELEASE FORM

NAME OF AUTHOR: Richard Ballard Fletcher

TITLE OF THESIS: Body Composition and Self-Esteem: A Validation of Selected Self-Esteem Measures

DEGREE: Master of Science

YEAR THIS DEGREE WAS GRANTED: 1995

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly, or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, except as hereinbefore provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's written permission.

R.B Fletcher

1302 Galbraith House

Michener Park,

Edmonton

Alberta, T6H 5B5

DATED: 27 September, 1995

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read and, recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled **Body Composition and Self-Esteem: A Validation of Selected Self-Esteem Measures** submitted by **Richard Ballard Fletcher** in partial fulfillment of the requirements for the degree of **Master of Science.**

Dr. J D Marshall
(Supervisor)

Dr. A B Nielsen
(Supervisor)

Dr. J Hogg
(Committee Member)

Dr. T O Maguire
(Committee Member)

DATED SEPT. 27/95

DEDICATION

To Sue for all her support and encouragement, and to my parents Joyce and my late father George, thank you all for everything.

ABSTRACT

The main purpose of this study was to determine whether three self-esteem measures based on three theoretical models of self-esteem classified subjects in a similar pattern. The three measures were the Rosenberg Self-Esteem Scale,(RSES) (Rosenberg, 1979), Piers-Harris Children's Self-Concept Scale, (PHCSCS)(Piers, 1984),and SDQ II (SDQ) (Marsh, 1990a). A secondary purpose was to determine the relationship between body composition and self-esteem by examining where total and subscale self-esteem scores differed in obese adolescent males and females.

322 adolescent males and females were administered the RSES, PHCSCS and SDQ self-esteem questionnaires. Height, weight and sum of five skinfolds (triceps, biceps, subscapular, suprailliac, medial calf) (CSTF, 1986) were obtained for each subject. Obese males (n = 27) and females (n = 36) were age, gender and school matched with a control to form a subgroup of n = 126.

Loglinear analysis suggested conflicting patterns of relationships between the three self-esteem measures. Results further indicated the RSES to be the measure contributing most to the poor patterns of relationships among the measures.

Overall, girls reported lower mean total scores on all three self-esteem measures. Furthermore, girls reported lower self-esteem than boys on the PHCSCS subscales for anxiety, physical attributes and appearance, and happiness. Also, results on the SDQ indicated girls to be lower on the physical appearance, physical ability, general self, emotion, and opposite sex relationships subscales. Obese compared to nonobese subjects reported lower mean total RSES and SDQ scores. Obese compared to nonobese reported

lower scores on the PHCSCS physical, popularity and happiness subscales and SDQ physical appearance, physical ability, general self and same sex relationships subscales. Obese girls compared to obese boys and nonobese girls reported lower mean total scores for the RSES and the SDQ, and also on the PHCSCS anxiety subscales and SDQ general self and physical ability subscales.

Results suggested that girls have lower self-esteem both at global and certain subscale levels when compared to boys. Furthermore, it was obese girls who reported lower self-esteem scores both globally and in certain subscales when compared to nonobese girls and obese boys. The differences found, suggested that data for males and females should be analyzed separately when examining relationships between obesity and self-esteem.

ACKNOWLEDGEMENTS

I would like to thank my supervisors Dr. Dru Marshall and Dr. Brian Nielsen for all their support and encouragement throughout my program. You have both provided a positive learning experience which will continue to influence me in the years to come. I owe much to you both. I would also like to thank Dr. Tom Maguire for his excellent teaching and eternal patience in furthering my understanding in statistics. Also thanks go to Dr. John Hogg, whose insightful comments have sharpened my knowledge of the self-esteem area.

Special thanks go to Lara Bloxham who unselfishly gave so much of her time helping to collect the data. I owe you big time Lara. Thanks also to Catherine, Helene, Julie and Mark for helping to record data. I would also like to acknowledge the special friendship and help given by Tony Webster in both collecting data and drinking coffee. Cheers Tony !

Thanks also to the principals of the following Junior High Schools: Ottewell, Highlands, Hardisty, Lawton, Westlawn, Rossyln and Vernon Barford. Above all, I would like to say a big thank you to all the students who so willingly took part in the research and who made the whole thing possible.

CONTENTS

Chapter		Page
1	**INTRODUCTION**	1
	Purpose	4
	Limitations	4
	Delimitations	4
	Definitions	5
2	**REVIEW OF THE LITERATURE**	8
	Introduction	8
	Part I	9
	"Self" as a Concept	9
	William James's Conception of Self	9
	Charles H Cooley and the Social Emergence of "Self"	10
	George H Mead and the Emergence of "Self" through Symbolic Interaction	11
	Self-Esteem Theories	12
	Summary	15
	Theoretical Models of Self-Esteem	15
	Part II	17
	Pediatric Obesity and Self-Esteem	17
	Obesity Classification	19
	Self-Esteem and Obesity	20

Part III 26

Validity 26

Problems Associated with Measuring

Self-Referent Variables 29

Construct Validation Methods 29

Multitrait-Multimethod Analysis 30

Factor Analysis 31

Self-Concept/Esteem Validation 32

Conclusion 42

3. **METHODS AND PROCEDURES** 45

Subjects 45

Measures 45

Anthropometric Measures 45

Classification of Obesity 46

Self-Esteem Measurement 46

Self-Esteem Measures 47

Statistical Analyses 48

4. **RESULTS AND DISCUSSION** 50

Subject Characteristics 50

Reliability 52

Discussion of Reliability 55

Validity 57

Discussion of Validity 60

MANOVA 64

Discussion of MANOVA Results 71

5. **SUMMARY AND CONCLUSIONS** 76

Suggested Future Research 78

Conclusion 80

6. **REFERENCES** 82

7. **APPENDICES** 88

Appendix A 88

Consent Form 89

Appendix B 92

Rosenberg Self-Esteem Scale 93

Piers-Harris Children's Self-Concept Scale 94

Self-Description Questionnaire 95

LIST OF TABLES

Table		Page
1.	Studies Analyzing the Relationship Between Pediatric Obesity and Self-Esteem.	22
2.	Group Descriptives for Gender.	50
3.	Descriptives for Matched Obese and Nonobese Subjects.	51
4.	Internal Consistency and Standard Errors of Measurement for RSES, PHCSCS and SDQ.	53
5.	Internal Consistency and Test-Retest Total Score for RSES and Total and Subscale Scores for PHCSCS and SDQ.	55
6.	Chi-Square Loglinear Analysis with Backwards Elimination for RSES, PHCSCS and SDQ.	58
7.	Cross Tabulation of Agreement Between SDQ and PHCSCS.	59
8.	Cross Tabulation of Agreement Between PHCSCS and RSES.	59
9.	Cross Tabulation of Agreement Between SDQ and RSES.	59
10.	Spearman Rank Order Correlations Between RSES, PHCSCS and SDQ.	60
11.	Means and Standard Deviations for Total RSES Score and Total and Subscale Scores for the PHCSCS and the SDQ For Boys and Girls.	65
12.	Means and Standard Deviations for RSES Total Score and PHCSCS and SDQ Total and Subscale Scores for Matched Obese and Nonobese Subjects.	68

13. Means and Standard Deviations for Total RSES Score and PHCSCS and SDQ Total and Subscale Scores for Matched Obese Girls and Obese Boys 69

14. Means and Standard Deviations for Total RSES Scores and PHCSCS and SDQ Total and Subscale Scores for Matched Obese Girls and Nonobese Girls 70

CHAPTER 1

Introduction

> "Self-esteem is nearly as ubiquitous a construct as intelligence, but there is less agreement about how to measure it." (Blascovich & Tomaka, 1991) (p-116)

The terms self-concept and self-esteem have often been used interchangeably. The lack of distinction between the constructs is due to the overlap that exists between self-concept and self-esteem. Self-concept is commonly accepted as the descriptive aspects that are used to define *self,* whereas self-esteem is the evaluative component of self-concept. Many instruments purport to measure self-concept, however, closer examination of the wording indicates a strong evaluative theme underpinning most instruments. Thus, although the distinction between self-concept and self-esteem can be made, operationalizing a measure specific to each of the constructs is difficult (Hattie & Marsh, in press).

Self-esteem is an important variable in the psychological well being of an individual. In order to gain an accurate understanding of how individuals perceive themselves one needs instruments that are theoretically and psychometrically sound. Surprisingly few studies have assessed the construct validity of the vast array of self-esteem measures that are available (Byrne, 1983; Hattie, 1992; Wells & Marwell, 1976; Wylie, 1974). Studies that have addressed the construct validity of self-esteem have employed a variety of approaches, from correlational (Demo & Savin-Williams, 1983; Savin-Williams & Jaquish, 1981) and multi-trait multi-method (Hamilton 1971; Silber &

Tippett, 1965; Winnie, Marx & Taylor, 1977) to factor analysis (Byrne, 1989; Demo, 1985; Fleming & Courtney, 1984; Marsh, 1990b; Marsh & Smith, 1982; Shavelson & Bolus 1982; Van-Tuinen & Ramanaiah 1979). The primary aim of these studies has been to assess the degree of association between instruments purporting to measure the same construct. No known study has assessed whether or not self-esteem measures derived from different theoretical models classify subjects in similar patterns. From a logical perspective, the functional worth of a measure must be demonstrated through its ability to provide similar patterns of results with other instruments purporting to measure the same construct.

The dearth of self-esteem construct validation literature has been confounded by the lack of theoretical consensus as to how perceptions of self are organized and measured. Three theoretical models: unidimensional, multidimensional, and hierarchical, have been proposed to represent the nature of self-esteem. As will be shown later various measurement instruments have been designed based on these theoretical models. Evidence is available for construct validity within unidimensional (Demo, 1985; Demo & Savin-Williams, 1983; Hamilton, 1971; Savin-Williams & Jaquish, 1981; Silber & Tippett, 1965) and multidimensional measures of self-esteem (Shavelson & Bolus, 1982; Winnie & Marx, 1977). There are no known studies examining the convergent validity of self-esteem measures based across each of the three theoretical models of self-esteem. Consequently the relationship between self-esteem and other indicators in children is often controversial . The controversy may in part be due to a measurement problem with self-esteem itself, since it is possible that the competing models of self-esteem may in fact be

measuring different parameters. Thus it is important to establish whether three self-esteem instruments based on the various theoretical models do classify subjects in a similar pattern as they all claim to assess self-esteem.

Various levels of self-esteem have been reported for different subgroups of children. For example, the relationship between childhood obesity and self-esteem is unclear, with some studies indicating self-esteem to be lower in obese children (Corbin, Pangrazi, Corbin, Peterson, Pangrazzi & Meza, 1995; Drake, 1988; Klesges, Haddock, Stein, Klesges, Eck & Hanson 1992; Manus & Killeen 1995; Martin, Housley, McCoy et al, 1988; Sallade, 1972; Strauss, Smith, Frame & Forehand, 1985) and others suggesting self-esteem in obese children to be similar to nonobese peers (Kaplan & Wadden 1986; Kimm, Sweeney, Janosky & MacMillan, 1991; Mendolson & White 1982; O'Brien, Smith, Bush & Peleg, 1990; Wadden, Foster, Brownell & Finley, 1984). Establishing whether self-esteem instruments based on the three different theoretical models of self-esteem classify subjects in a similar manner should provide a framework from which to evaluate the inconsistencies in the pediatric obesity self-esteem literature.

The purpose of this study was to provide a cross theoretical model analysis of three self-esteem measures to determine whether or not they classify subjects in consistent patterns and provide evidence to explain the inconsistancies in the literature. Also, analyzing more advanced theoretical models of self-esteem should help to indicate, using the subscale scores, where self-esteem scores differ in obese and nonobese children. Thus, an indication of which model provides the most sensitive assessment of self-esteem in obese children should emerge, potentially providing more insightful information as to the

types of intervention that may be derived. This study is intended to clarify these issues. Although two of the measures used in this study claim to assess self-concept (Piers-Harris Children's Self-Concept Scale (PHCSCS) Piers, 1984; Self-Description Questionnaire II (SDQ) Marsh, 1990 a), the evaluative wording of items is more indicative of self-esteem. Thus, for this study, self-esteem is the construct employed and all results obtained with the three instruments are considered to be assessments of this construct.

Purpose of the Study

The primary purpose of the study was to assess whether three self-esteem measures: 1. Rosenberg Self-Esteem Scale (RSES) (Rosenberg, 1979) (Unidimensional) , 2. PHCSCS (Piers, 1984) (Multidimensional), and 3. SDQ (Marsh, 1990 a) (Hierarchical), derived from three different theoretical models of self-esteem classify subjects in a similar manner. It was expected that obese children would exhibit lower self-esteem scores on certain subscales than their nonobese peers on the PHCSCS and the SDQ. Thus a secondary purpose was to determine whether or not differences exist in self-esteem between obese and nonobese adolescents on total and subscale self-esteem scores.

Limitations

1. Subjects were drawn from schools offering different programs so the effect of these programs was not controlled for and therefore may have affected self-esteem scores.

Delimitations

1. Questionnaires used (See Appendix B) were: 1. RSES, (Rosenberg, 1979), 2.

PHCSCS, (Piers, 1984) and 3. SDQ, (Marsh, 1990 a). Selection of questionnaires was based on: 1. The number of citations per year in research articles (Blascvich & Tomaka 1991), and 2. their suitability for use with adolescent males and females in grades 8 and 9.

2. No interviews were conducted.

3. As only adolescent males and females (Grades 8 and 9) were used, the results from this study are only generalizable to this age range.

4. Only Junior High Public Schools in the Edmonton area were chosen. Private schools and schools for the handicapped were not accessed.

5. Only one measure to define obesity was employed: a sum of five skinfolds (CSTF, 1986).

Definitions

Self-Concept. Self-concept is comprised of the descriptive aspects used to define self and was defined by Shavelson, Hubner and Stanton (1976) as an individual's perceptions of him/herself, which are formed through their interaction with the environment. These perceptions are influenced either by significant others or environmental reinforcements. However, these perceptions are thought to influence a person's actions, which ultimately affect the perception of one's self.

Self-Esteem. Sel-esteem is the evaluative component of self-concept. Gruber (1986) stated that " Self-esteem is the value we place on our self-image" (pp-30). As Fox (1992) specified, self-esteem ".. carries an evaluative global judgement of self on whatever criteria the individual in question considers important." However, global self-esteem may

not be specific, in that content of what is important will vary from person to person. Blascovich and Tomaka (1991) state: "...self-esteem is a hypothetical construct that is quantified for example as the sum of evaluations across salient attributes of one's self or personality. It is the overall affective evaluation of one's own worth, value or importance." (p. 115).

Unidimensional Self-Concept. The assumption that underlies the unidimensional model of self-concept is that the various domains of self are encapsulated into a single entity. As a result self-concept is viewed as being global with each domain contributing equally to the overall self-perception. Although various aspects of "self" are measured, this model does not account for more specific areas of low/high self-esteem, given that a total, rather than the domain score is used as an indication of an individual's self-esteem.

Multidimensional Self-Concept. This model seeks to identify more specific self-perceptions enabling more facets to be included in measurement scales. Multidimensionality views global "self" as being central with subdomains contributing to the overall self-perception. The identification of subdomains permits identification of areas that contribute to low/high self-esteem.

Hierarchical Self-Concept. This model is hierarchically constructed with global self at the apex and more generalized domain self-concepts (eg, physical, academic, social) at the next level, and even further specific subdomains (eg, physical competence, physical attractiveness, strength) at the final level. Thus specific subdomains act to influence the domains which in turn affect global self.

Obesity: Obesity is defined as an excess of body fat leading to a significant

impairment in health (Burton, Foster, Hirsch & Van Itallie, 1985). For the purposes of this study, obesity was operationally defined as being greater than or equal to the 85th percentile of Canadian population norms (Canada Fitness Survey, 1985) for the sum of five skinfolds (triceps, biceps, subscapular, iliac crest, medial calf) (CSTF, 1986).

CHAPTER 2

Review of Literature

Introduction

The relationship between pediatric obesity and self-esteem is unclear, with some studies indicating lower self-esteem in obese children (Corbin et al 1995; Drake 1988; Klesges et al 1992; Manus & Killeen 1995; Martin et al 1988; Sallade 1972; Strauss et al, 1985) and others demonstrating comparable self-esteem in obese children to their leaner counterparts (Kaplan & Wadden 1986; Kimm et al 1991; Mendolson & White 1982; O'Brien et al, 1990; Wadden, et al 1984). This review aims to address, from a theoretical perspective, some potential reasons for such diversity in results.

Section one focuses on the phenomenon of "self" as a philosophical/psychological construct is clarified with attention to the contexts in which it develops. Because self-concept and self-esteem have been used interchangeably it is important to make the distinction between the two terms in order to alleviate confusion. Relevant theories of self-esteem are reviewed to highlight the research that underscores self-esteem measurement. Finally three theoretical models of self-esteem are examined critically to ascertain their relative merits and weaknesses.

In the second section the pertinent pediatric obesity and self-esteem literature is analyzed using the theoretical models of self-esteem. Due to the diversity of instruments used to measure both self-esteem and obesity this section provides an overview of the limitations of establishing these two parameters in this population.

In the final section the notion of validity is addressed both from a general and an

applied perspective. General interpretations of validity are discussed in order to highlight the theoretical and methodological approaches to validating constructs. Furthermore, self-concept/esteem validation studies are reviewed in order to establish construct validity.

Part I

"Self" as a Concept

"Self" is a highly ambiguous expression that has been associated with a variety of phenomena such as self-concept, self-esteem, self-image, self-attitude, self-efficacy and self-worth. Wells and Marwell (1976) suggested that:

> "All of these terms denote some basic process of psychological functioning which can be described as either self evaluation or self-affection or a combination of the two." (p. 8)

When using the term "self" one needs to ascertain the psychological function to which it refers. Analysis of the term "self" should illuminate our understanding of why it has been used in such an eclectic manner.

William James's Conception of "Self"

The psychologist William James (1890) viewed "self" as conscious phenomenon founded upon thoughts, objects and the experiences. Self was dichotomized into the "me" and the "I", with the "me" representing the self as the known, and the "I" representing the self as the agent of experience. The "me" was dimensionalized into the material, social and spiritual. The material "me" encompassed the body, belongings and family and was hierarchically organized with some aspects weighted with different degrees of intimacy.

The social "me" was that which received recognition from other people. As James (1890) suggested... "Properly speaking a man has as many social selves as there are individuals who recognised him, and carry an image of him in their mind." (p. 249). Therefore the social "self" was constructed through others perceptions of one's self.

The spiritual "me" was based upon emotions and feelings. Addressing the affective side of self presupposed the ability to make self-evaluations based upon feelings and emotions and how these were affected by individual experiences. James (1890) argued self-esteem was the ratio of one's success to one's aspirations. Success therefore influenced self-esteem which was equated to a barometer in which esteem was either raised or lowered according to the situation.

James (1890) suggested the "I", or the pure ego, was the agent of experience which represented the "self" as the thinker. Because thoughts and objects were based on conscious experience, the relationship between the "I" and the "me" became essential for a conscious "self" to exist. In sum the "I" placed the known "me" into a conscious self based on experience.

Charles H Cooley and the Social Emergence of "Self"

Cooley (1902) also viewed self as a conscious phenomena emphasizing James's (1890) social dimension Self, it was argued, stemmed from an individual's interaction with society. Cooley (1902) contended self-awareness was an interaction of three states of consciousness: self consciousness (what the individual thinks of him/herself), social consciousness (what the individual thinks of other people) and public consciousness (a collective view of previous states of consciousness organized into a communicative

group). Thus Cooley (1902) proposed the "looking glass self" in which the individual perceived how other people perceive him/herself. To what degree self-feeling was affected was determined by the value placed on those chosen to view oneself from. Potentially the individual formed a sense of positive "self" based on perceptions of those they perceived valued them and dispensed with those they perceived did not. However, this was all a part of the larger social arena in which social interaction aided the confirmation or the reinterpretation of others' perceptions of one's self.

George H Mead and the Emergence of "Self" through Symbolic Interaction

Mead (1934) proposed a developmental process in which both the "mind" and "self" arise through social interaction. The acquisition of a "mind" was a prerequisite for self development. Mead (1934) suggested that gestures, and in particular the vocal gesture, was the mechanism by which "mind" and "self" appeared. Language provided conscious communication which had both significance to the individual and to others. As a result common meaning was attached to gestures and words enabling a mind to be acquired which in turn formed a sense of self based on social communication.

Language further permitted the individual to become an object to him/herself through the transfer of others attitudes and views which then formed a core of self-conscious behaviour. As a result, self-attitudes arose from social communication to form a central component in "self" development. This hypothesis was later tested by Rosenberg (1979) who noted that with age children learn through social communication and interaction to evaluate others' opinions and attitudes by accepting those they value and dispensing with those they do not.

Because evaluation was a dimension of attitudes one was likely to compare oneself to others or past experiences. As a consequence positive/negative self-evaluation could likely affect self-esteem. Mead (1934) further proposed the play and the game stage to explain the genesis of the reflective "self". Through play children view themselves from the viewpoint of others by taking on different roles. The game stage, however, involved the attitudes of others and therefore the child had to understand the structure of the game and coordinate the perspectives of all involved. Thus from the game stage emerges an organized personality that is conscious of "self" in relation to others.

Self-Esteem Theories

Building on the work of Cooley (1902) and Mead (1934), Rosenberg (1979) provided a more in-depth analysis of self-esteem. Rosenberg (1979) provided a developmental approach to studying adolescent self-esteem. Central was the view of self-image as being global, encompassing all evaluations of "self" in the various domains and experiences based on self-estimation. As Wells and Marwell (1976) suggested... " With each self-estimate weighted by its corresponding self-value, the overall self-esteem of the individual represents some kind of psychological summation of these specific weighted self-evaluations." (p. 31). This complemented James's (1890), view that some aspects of the "me" may be more intimate than others with the individual valuing more highly those aspects they perceive as being of more worth than others that are less positive.

Furthering Mead's (1934) work, Rosenberg (1979) observed that younger children tended to be unreflective in their thoughts and based their self-concept around personal, social and general characteristics. Because they do not view themselves from the

perspective of others, self-concept is mainly derived from adult feedback and is accepted at face value. The focus was therefore more on outward discovery, than inward analysis. As a consequence, Rosenberg (1979) stated " ..self-consciousness is generally low, self-concept stability high and self-esteem satisfactory". (p. 245) Rosenberg (1979) further proposed that with an increase in age children developed internal perceptions of self by becoming more aware of their feelings, thoughts and experiences. Moreover, a sense of self-consciousness emerged based on how other people may perceive them, fostering either a positive or negative perception of self. As Rosenberg (1979) suggested ... " What were formally unquestioned self-truths now become problematic self-hypothesis and the search for the truth is on." (p. 255) This shift towards an introspective analysis is formed around perceptions of how others view oneself. Consequently, self-concept becomes more volatile and subject to influences and interpretations that were previously unquestioned. As a result self-esteem can be more readily affected.

For Coopersmith (1959) " self-esteem is a personal judgement of one's worthiness that is expressed in the attitudes the individual holds towards himself." (p. 5) Moreover, self-esteem was derived from two perspectives. The first, subjective expression, was centred around the individuals' self-perception and description. The second, behavioral expression, was a behavioral manifestation of an individual's self-esteem that can be observed by others. A further distinction was made between genuine and defensive self-esteem. Genuine self-esteem was based on worthiness and expression of success, whilst defensive self-esteem was a denial of subjective experience.

Coopersmith (1967) identified five types of self-esteem based on subjective self-

evaluation and observed behavioral ratings. Those high on both the subjective self-evaluation and behavioral rating were more academically and socially content. The high-lows (subjective-behavioral) tended to be those who maintained a positive self-image but displayed low behavioral congruence. The medium-medium's were considered to be the typical average child with moderate aptitude and attainment. The low-highs were those who had negative self-evaluations but demonstrated high behavioral self-esteem. These children were more anxious with higher levels of motivation and judged themselves by their own personal criteria rather than that of others. The low-lows tended to be less social and academic and accepted the reality of their situation. Establishing congruence between self-evaluation and behavioral expression provided an indication of an individuals true or false self-esteem. Coopersmith (1967) thus made the distinction between the experienced self (subjective self-report) and presented self (behavioral ratings of self-esteem observed by others).

It was further contended by Coopersmith (1967) that self-evaluations were underscored by success, values, aspirations and defense, as he suggested:

> "The process of self-judgement derives from a subjective judgement of success, with that appraisal weighted according to the value placed upon different areas of capacity and performance measured against a person's personal goals and standards filtered through his capacity to defend against presumed or actual occurrences of failure." (p.242)

Coopersmith's (1967) view was similar to James' (1890) in that self-evaluation of one's aspirations was driven by perceptions of success or failure with the appropriate value ascribed to these and that the ratio between these was an indication of a person's overall

self-esteem.

Summary

For James (1890), Cooley (1902) and Mead, (1934) self-concept was an introspective view of oneself based on personal experience and interactions with the environment. Furthermore, perceptions of self could be evaluated against the experiences to provide an indication of one's relative self-esteem. Thus self-esteem was the evaluative component of self-concept.

Self-esteem was analyzed in more depth by Rosenberg (1979) and Coopersmith (1967) who provided a developmental framework from which theoretical self-concept/esteem models could emerge and be empirically tested.

Theoretical Models of Self-Esteem

The theoretical disparities as to how self-esteem is constructed have led to the development of three conceptual models. The first is a unidimensional model of self-esteem. Measurement as set out by Coopersmith (1967) and Rosenberg (1979) proposed self-esteem as a single unit entity comprised of various aspects of self, such as social, academic and physical self-esteem. Moreover, each aspect of self carried a subjective weighting with the total point score being indicative of one's global self-esteem. However, as Fox (1988) observed simply adding the scores on the various aspects, such as the academic, social and physical self, fails to acknowledge the underlying mechanisms of self-esteem. Consequently a total score may not be truly indicative of how an individual feels about him/herself. Unidimensionality fails to acknowledge the individuality of self-esteem dimensions and does not address the specificity of certain situations, values and beliefs of

each individual. As a result total scores classify subjects as having either high or low self-esteem with no attention paid to where low/high areas may be manifest. As an initial step in measurement, this model provided an adequate indication of ones relative self-esteem and its utility is better suited as a basic screening device for identifying low/high general self-esteem .

The second model was based on a multidimensional perspective as proposed by Shavelson et al (1976). This model seeks to identify more specific self-perceptions enabling more facets to be included in measurement scales. Multidimensionality views global self as being central with various dimensions of self contributing to the overall self-perception. The benefit of this is that one can identify the individual dimensions that contribute to low/high self-esteem. Multidimensionality has the advantage over unidimensionality in that it allows for more exploration of the underlying domains and for examination of the relationships between variables (Fox & Corbin , 1989). Furthermore it allows for appropriately targeted interventions, directed towards domains that indicate low self-esteem, such as self-esteem enhancement programs.

Finally, the hierarchical construct of self is also based on the work of Shavelson et al (1976) later confirmed by Shavelson & Bolus (1982) and modified by Marsh & Shavelson (1985) and Marsh (1987). The model is hierarchically organised with global self at the apex. Blascovich & Tomaka (1991) noted that:

> " The facet model incorporates a generalized sense of self, more specific facets or domain self-concepts and even more specific skills and abilities. In short, specific skills and abilities contribute to facet selves, which in turn contribute to general sense of self." (p. 154)

The hierarchical model allows a more in depth approach to understanding how self-esteem is constructed and how the subdomains affect the higher levels. It further allows individual domains to be assessed and for examination of relationships that may exist between subdomains.

Self-esteem measures derived from multidimensional and hierarchical models have the advantage of being able to detect finer differences between and within individuals. They reflect components at a different level and yield greater information as to how self-concept is constructed. Self-esteem measures utilizing multidimensional and hierarchical models have greater practical applications as they allow for intervention to be directed towards the areas that indicate low self-esteem. Multidimensionality provides what Fox (1988) described as a more personalized approach to understanding self-esteem structure, rendering scope for the individual to individualize and customize their self-esteem structure. The individual is therefore able to attach more weight to areas of personal importance and value and is thus able to self-serve their self-esteem.

Part II.

Pediatric Obesity and Self-Esteem

Obesity is a multifaceted disorder with an increasing prevalence (Gortmaker, Dietz, Sobol & Wehler, 1987) and is defined as an excess of body fat leading to a significant impairment in health (Burton, Foster, Hirsch & Van Itallie, 1985). Williams (1986) suggested that etiology of obesity is due to endogenous and exogenous variables:

> "Endogenous factors include genetics, neuroendocrine imbalances and metabolic disturbances whereas exogenous factors involve socioeconomic status, cultural background, nutrition, physical activity and emotional states." (p. 9)

Health related problems associated with adult obesity include coronary heart disease, elevated blood pressure and orthopaedic problems. However, these health problems are not always evident in children. Of more importance in obese children are the psychosocial problems which include rejection by peers, depression, low self-concept, (Strauss et al, 1985), social stigmatization and prejudice (Allon, 1979; Lerner & Gellert, 1969; Wadden & Stunkard, 1985). Behavioral problems include social withdrawal and lack of social competence (Israel & Shapiro, 1985)

Considering the negative psychosocial problems reported in obese children one might conclude that this population has low self-esteem. However, evidence tends to be equivocal. Some research has indicated low self-esteem in obese children compared to their nonobese peers (Corbin et al 1995; Drake, 1988; Manus & Killeen 1995; Martin et al, 1988; Sallade, 1972; Strauss et al, 1985), whilst other research found no differences in self-esteem between obese and nonobese children (Kaplan & Wadden, 1986; Kimm et al, 1991; Mendolson & White, 1982; O'Brien 1990; Wadden et al, 1984) (See table 1). Clarifying this issue is important since self-esteem may be a mediating factor in the onset, maintenance or reduction in obesity.

The four main sources of the discrepancies in the pediatric obesity and self-esteem literature are: 1) Problems associated with the measurement of body composition; 2) The cut off point for classifying levels of obesity (Garn, Leonard & Hawthorne, 1986; Lohman

1992; Marshall, Hazlett, Spady, & Quinney, 1990; Marshall, Hazlett, Spady, Conger & Quinney, 1991; McLaren, 1987); 3) The theoretical models used in the design of measurement instruments for self-esteem; 4) The lack of a gold standard by which to measure self-esteem,

Obesity Classification

A characteristic of most pediatric obesity and self-esteem studies has been the use of convenient field measures for estimating obesity (eg, height/weight tables). Recent developments in this field have led to a questioning of the validity of obesity classification generally, with some researchers outlining the major weaknesses of these field measures (Garn et al, 1986; Lohman, 1992; Marshall et al, 1990; Marshall, et al 1991; McLaren, 1987). When dealing with such an important issue as body composition in obese children it is essential that the accuracy of one's estimation of body fat is valid. Of equal importance is establishing where to place the cut-off to define health risks and thus obesity. Since overt health risks are not readily apparent in obese children and adolescents, then obesity classifications tend to be statistical value laden judgements. As a consequence incorrect classifications may result, leading to flawed data which in the long term will create confusion and result in misleading conclusions. Until a cut-off criterion is universally accepted, one may be incorrectly identifying individuals as obese. Field measures may therefore contribute towards the disparity in the literature given that height and weight scales have been so widely employed (Drake, 1988; Kaplan & Wadden, 1986; Kimm et al, 1991; Martin et al, 1988; Sallade, 1972; Wadden et al, 1984).

One must also question whether identification of obese children is related to self-

perception of obesity. If children do not perceive themselves to be obese then their size may not be an influential factor in determining their self-esteem. Alternatively self-perceptions of being overweight may have negative consequences on self-esteem regardless of body size. Identifying children as being obese may require supplementary evidence, such as a self-perception of obesity measure, in order to comprehend whether obesity positively or negatively impacts on self-esteem.

Self-Esteem and Obesity

The lack of consensus in the pediatric obesity and self-esteem literature stem at least in part from the lack of consensus as to how self-concept is constructed. The lack of agreement ultimately results in diverse outcomes when different instruments constructed under different theoretical models are used. Data will continue to be controversial until a 'gold standard' by which to measure self-esteem is reached.

Research that has employed unidimensional measures of self-esteem has provided a limited understanding of the relationship between self-esteem and childhood obesity (Martin et al, 1988; O'Brien et al, 1991). The inability of unidimensional measures to elicit subscales scores make it difficult to detect where areas of low self-esteem are manifest. The utility of unidimensional measures of self-esteem may be more suited only to detecting differences between obese and nonobese children. As a result data tend to be descriptive and have failed to address the underlying factors that may affect self-esteem.

Multidimensional measures have been used to analyze childhood obesity and self-esteem (Corbin et al 1995; Drake, 1988; Kaplan & Wadden, 1986; Kimm et al, 1991; Klesges et al 1992; Manus & Killeen 1995; Mendolson & White, 1982; Sallade, 1972;

Strauss et al, 1985; Wadden et al, 1984). Multidimensional measures have the potential to reveal particular areas where low self-esteem is evident. For example, Kimm et al (1991) noted subscale scores for physical appearance, popularity, anxiety, and happiness/satisfaction were lower in obese children than their scores for behaviour and intellectual school status. Manus & Killeen, (1995) also noted lower reported scores for obese children on general self-worth, social acceptance, physical appearance and athletic competence. Corbin et al (1995) reported that physical appearance was the dominant factor that discriminated between obese and nonobese boys. Klesges et al (1992), however, only reported physical self-competence decreased as body fat increased. These results indicate the usefulness of multidimensional measures, since the inclusion of subscales permits more of the underlying relationships to be revealed. This serves as a valuable tool for recognizing where differences in self-esteem subscales scores can be located within and between people, and helps to direct intervention appropriately.

Although multidimensional measures provide richer data they have demonstrated divergent results when the same scale is used, with those indicating self-esteem to be lower in obese children (Sallade, 1972, Strauss et al, 1985; Kimm et al, 1991; Klesges et al 1992; Corbin et al 1995; Manus & Killeen 1995) and those observing self-esteem to be similar between subjects of different levels of obesity (Mendolson & White, 1982; Wadden et al, 1984; Kaplan & Wadden, 1986; Drake, 1988). Consequently the relationship between childhood obesity and self-esteem is unclear even when using multidimensional measures. Although multidimensional scales have been widely used with obese children very few authors provide information on the subscale scores. Instead self-esteem is

reported as either being higher or lower with no reference to any subscale scores.

Table 1. Studies Analyzing the Relationship Between Pediatric Obesity and Self-Esteem

Author	Subjects	Measures	Results
Martin et al, 1988	n = 508 females Age 14-16	Rosenberg Self-Esteem Scale, 1965 The Quetelet Index Ht, Wt (m)2	Lower self-esteem as weight increased
O'Brien et al, 1990	n = 1003 black children 46 % males 54 % females Age 9-12 Subject either obese at Time 1 to 2, Obese at time 1 not at 2, Not obese at time 1 obese at time 2, Neither obese at time 1 or 2	Rosenberg Self-Esteem Scale, 1972 Obesity defined as 120% $\geq$ NHANES II, based on age, gender, Ht (measured on 2 occasions)	Overall sample self-esteem increased. Nonobese to nonobese increased self-esteem. Nonobese to obese, obese to obese, and obese to nonobese no change in self-esteem
Sallade, 1973	n = 240 Caucasian subjects 120 = obese 120 = control Age 8-16	Piers-Harris Children's Self-Concept Scale 1964 Ht, wt. for age norms, 75 th percentile of wt for age, sex on Revised Stuart Meridith Tables as cut-off for obesity	Lower self-esteem in obese children

Continued overleaf

Table 1. Continued

Author	Subjects	Measures	Results
Mendolson & White, 1982	n = 36 15 boys 21 girls 5 males obese 11 females obese Age 7-13	Piers-Harris Children's Self-Concept Scale 1964 Visual identification (U.S. Public Health Service, 1975), relative Wt for Ht placed at the 15 % overweight	No difference in self-esteem between obese and nonobese subjects
Wadden et al, 1984	n = 105 obese n = 210 nonobese 58 obese males 47 obese females All subjects were Caucasian Age 8-13	Piers-Harris Children's Self-Concept Scale 1969 > 20 % of ideal bodyweight for age, sex, Ht	No difference in self-esteem between obese and nonobese subjects
Strauss et al, 1985	n = 36 18 obese 18 nonobese 7 obese females 11 obese males Grades 2-5	Piers-Harris Children's Self-Concept Scale 1964 Visual rating of obesity, 1-5 scale. Three raters had to designate subject 4-5 to be obese	Lower overall self-esteem in obese subjects. Lower physical self-esteem for obese subjects

Continued overleaf

Table 1. Continued

Author	Subjects	Measures	Results
Kaplan & Wadden, 1986	n = 851 black children 49 % males 51 % females Age 9-18 Subjects divided into 4 groups based on relative BMI	Revised Piers-Harris Children's Self-Concept Scale 1984 The Queletet Index, expressed as % of ideal BMI using Ten State Nutritional Survey 1968-70. 50 th percentile Ht, Wt for age, sex considered to be ideal for calculating relative BMI	Self-esteem scores for obese subjects fell within normal range
Drake, 1988	n = 304 females 152 = obese 152 = controls Age 17-18	Revised Piers-Harris Children's Self-Concept Scale 1984 BMI expressed as a % of ideal BMI based on Ten State Nutritional Survey 1968-70. Obese ≥120 %	Self-esteem for obese subject lower than controls

Continued overleaf

Table 1. Continued

Author	Subjects	Measures	Results
Kimm et al, 1991	n = 130 21 black males 33 Caucasian males 38 black females 38 Caucasian females Age 8-17	Revised Piers-Harris Children's Self-Concept Scale 1984 Obesity based on % ideal Wt from National Centre for Health Statistics Ht, Wt tables. Moderately obese = 100-150 % Severely obese = 150 -200 % and Very severe > 200% than ideal Wt	Overall self-esteem scores for obese fell within norms. Boys reported lower levels of self-concept as weight increased on total scores.
Klesges et al (1992)	n = 132 31 obese males 17 obese females Age 3-5	Perceived Competence Scale for Children (1982) Obesity defined as 20 % body fat above ideal, based on national triceps skinfolds norms from U.S. Department of Health Education and Welfare (1977)	Lower physical self-competence as weight increased.

Continued Overleaf

Table 1. Continued

Author	Subjects	Measures	Results
Corbin et al (1995)	n = 523 males Grade 1-6	Martin Zaichowsky Self-Concept Scale (Grade 1-3) Harter Self-Perception Scale (Grade 3-6) National health-related fatness criteria. Combined subscapular and triceps skinfolds standard from the Fitnessgram National Test Battery.	Lower self-esteem for those above the national fatness criteria. Physical appearance the dominant factor that discriminated between obese and nonobese subjects.
Manus & Killeen (1995)	n = 45 18 obese 10 overweight 17 average weight Age 10-13	Self-Perception Profile Obese ≥ 20% median Wt for age Ht and sex on growth percentile charts	Obese subjects had lower self-concept than nonobese on global self worth, social acceptance, physical appearance and athletic competence subscale scores

Part III

Validity

Validity is a multifaceted and complex procedure. As Cook (1979) outlined:

> "Validity and invalidity refer to the best available approximation to the truth and falsity of propositions, including propositions about the cause..we should always use the modifier "approximately" when referring to validity, since one can never know what is true. At best one can know what has not yet been ruled out as false." (p. 37)

Since truth is rarely known, validation is never certain and is an ongoing process that adds plausibility to the construct under investigation. Messick (1989) noted "..validity is a matter of degree, not all or none." (p. 13) Furthermore, validity is enhanced/diminished as new evidence is accumulated. Therefore, at best, validation is only an estimate.

Cronbach & Meehl (1955) suggested the following types of validation: predictive, concurrent, content and construct related validity. However, Wells & Marwell (1976) pointed out that self-esteem is a private and unobservable phenomenon, at least not directly observable, rendering it difficult to ascertain from an individual's behaviour. Consequently predictive and concurrent validity have little application to self-esteem validation.

Content validity seeks to establish a sample of items that best represent the area under investigation. Moreover, it is how well the items fit the situations or subject matter on which the investigator bases conclusions. Content validity can therefore be incorporated into the process of construct validation as an initial stage in tapping the content of a proposed measure.

Cronbach & Meehl (1955) suggested a construct is a hypothetical trait of an individual or group of people that is reflected in a test performance. Because self-concept is an unobservable phenomena, Wylie (1974) argued that "Construct validity is necessary because by definition S's [subjects] cognitions and attitudes about himself..are private and beyond direct observation by the investigator." (p. 39). Construct validation is used when no criterion or valid indicator is available to confirm its presence and the investigator is

forced to construct and confirm their own criteria/indicator. Thus, the fit between one's conceptual and operational definition is important for construct validation. Furthermore, construct validation also provides evidence for the theory behind the measure as well as the measure itself.

The previous types of validity are seen as the traditional ways of what Messick (1989) terms the 'cutting of validity evidence'. This traditional approach, however, has been gradually replaced by a unitary view of validity in which all types of validity are subsumed under the title of construct related evidence. This new view of validity is as Rogers (1994) stated, one..." in which the distinction among types of validity gave way to varieties of evidence required in the validation process." Messick (1989) further proposed consequential validity, in which the utility, value implications and social consequences of the test score be included in the validation process. However, Maguire, Hattie & Haig (1994) have suggested that consequential validity should be applied to the arena of social debate rather than under construct validation process. Moreover, it should be the test users responsibility to address the consequences of the test use and score. The argument of consequential validity is as Rogers (1994) stated " not whether social consequences of a test use should be addressed, but rather who and where this assessment takes place."

In sum validity is a series of processes that when combined add to the plausibility of the result in hand. The more evidence accumulated the stronger the claim as to the validity of the measure or results obtained.

Problems Associated with Measuring Self-Referent Variables

Establishing construct validity is problematic due to external detracting factors. Smith & Glass (1987) suggested that the following characteristics could influence construct validity in self-report measures:

1. **Reactivity;** Subjects may be tempted to misrepresent their true characteristics.
2. **Insensitivity;** The failure of an indicator to reveal statistical significance, when treatment has been effective.
3. **Response sets;**
    i) *Social desirability;* Subjects respond in common rather than honest ways.
    ii) *Acquiesence;* Subjects agree with statements and questions.
    iii) *Extremity;* Subjects respond only to the highest and lowest alternative responses.
    iv) *Evasiveness;* Subjects tend to offer no opinion by opting for the middle alternative.
    v) *Carelessness;* Subjects answer thoughtlessly

Unless the above factors are considered and controlled for, serious threats to validity will result.

Construct Validation Methods

Two procedures that can be used to validate constructs are both correlational techniques. The first is the multitrait-multimethod analysis as outlined by Campbell and

Fiske (1959). The second technique is factor analysis.

Multitrait-Multimethod Analysis

Campbell and Fiske (1959) argued that evidence for construct validity is enhanced by establishing convergent and discriminant validity. As such, two independent methods that purport to measure the same multiple traits should correlate highly (convergent validity) and correlate negligibly with theoretically unrelated constructs (discriminant validity).

Supporting convergent validity, correlations between different methods purporting to tap the same trait would be high. These are called MONOTRAIT-HETEROMETHOD CORRELATIONS. Discriminant validity requires MONOTRAIT-HETEROMETHOD CORRELATIONS to exceed HETEROTRAIT-HETEROMETHOD CORRELATIONS and HETEROTRAIT-MONOMETHOD CORRELATIONS. Furthermore patterns of correlations between the various traits should be similar in both the HETEROMETHOD and MONOMETHOD blocks.

The multimethod-multitrait method is based on the assumption that score variance is divided into method and trait variance. Method variance pertains to specific aspects of the measurement instrument, while trait variance accounts for individual differences in the trait being examined. Using diverse methods that purport to measure the same construct avoids confounding method variance otherwise correlations may be explained by shared method variance. As Smith & Glass (1987) pointed out, "If the indicators not sharing method variance correlate positively with each other there is strong evidence for validity." (p. 108). Given the rigour of the multitrait-multimethod procedure one must ensure that

methods, not measures, are different. As self-esteem measurement is usually based on paper and pencil self-report, one might find that measures converge due to their similarity or mono-operational bias as outlined by Cook and Campbell (1979). As Messick (1989) observed;

> "..multiple measures or exemplars of a construct all based on the same procedural method, such ratings or a paper and pencil test, might still be grossly misleading due to the shared method variance." (p. 35)

However, since self-esteem is an introspective evaluation of oneself, one is limited in the variety of methods that can be adopted. Given that self report has been the common method used, establishing alternative valid methods are problematic. Attempts at validating different methods of measuring self-esteem will be highlighted in a later section.

Factor Analysis

Factor analysis is a statistical technique used to reduce large amounts of data into manageable subsets. It is applied to a single set of variables which then can be formed into relatively independent subsets. Groups of variables that positively correlate and are considered independent of the other subsets are thought to represent underlying constructs and are termed factors.

The process of factor extraction involves identifying variables that share common variance which are then grouped into subsets (factors). A second correlation is then used to identify relationships between the variables from the original matrix and each of the factors (factor loading). Variables should factor load highly within each factor, but not with other factors. Thus common variance of variables within each factor is established as

being relatively independent of other factors. Finally and if necessary, factors are mathematically rotated to enhance the variables correlations within each factor and to minimize the correlations with other factors.

Factor analysis therefore lends itself to establishing construct validity. Smith & Glass (1987), advocated that factor analysis "..detects whether all the variables share common variance....,or whether the variance is unique from trait to trait." (p. 221). Unlike the multitrait-multimethod matrix, factor analysis allows similar measures to be used in order to establish construct validity, as common variance is extracted. However, one must be aware that factor interpretations are, as Messick (1989) noted, hypotheses which themselves need further empirical testing. Factor analysis therefore offers an attractive alternative to the multitrait-multimethod technique when validating constructs.

Self-Concept/Esteem Validation

Studies that have addressed the construct validity of self-concept/esteem can be classified into those that have used: 1. Correlational analysis; 2. Multitrait-multimethod analysis; 3. Factor analysis; and 4. A combination of Multitrait-multimethod and factor analysis.

Savin-Williams and Jaquish (1981) analyzed measures of experienced and presented self-esteem based on traditional self-report (Rosenberg Self-Esteem Scale, Lerner Self-Description Scale),a situational self-report measure of self-esteem (Beeper Self-Reports as outlined by Savin-Williams & Jaquish, 1981), behavioral observation and peer rating of self-esteem (as outlined by Savin-Williams and Jaquish, 1981). Subjects, (n = 40), included 20 males and 20 females aged 12-14. Significant correlations (r = .72)

between Rosenberg's Self-Esteem Scale and the Lerner Self-Description Scale were observed but neither were highly correlated to the observational (r =.24 and -.20) and peer ratings of self-esteem (r =.27 and -.21). The Beeper Self-Report was not significantly correlated to observational (r =.01) and peer ratings of self-esteem (r = -.17). No correlation was given between Beeper Self Report and traditional measures. The evidence suggested no relationship between experienced and presented self-esteem. As self-reported feelings of self-esteem did not converge with projected self-esteem, it would appear that two distinct concepts were being measured, indicating no congruence with Coopersmith's (1967) proposed measurement of self-esteem.

Demo and Savin-Williams (1983) investigated the relationship between self-esteem and social class with father's occupation used to define social class. It was argued that self-esteem was a function of social class based on the theoretical assumptions of reflected appraisals, self-perception, social comparison, and psychological centrality. The aim of the study was to asses the relationship between social class and self-esteem with an increase in age. A sample of 850 subjects, included n = 48 % males, n = 52 % female from grades 5-8. Results indicated a stronger relationship between social class and self-esteem in eighth but not fifth graders. Self-esteem also increased from grade 5 to 8 indicating a developmental trend in self-acceptance. A feature of the study was the correlational data obtained on three self-report measures of self-esteem (Coopersmith Self-Esteem Inventory, Morolla Looking-glass Self-Esteem Inventory, Waetjen-Liddle Learner's Self-Concept Scale). Results indicated low to moderate correlations between instruments (r =.36 through .44). Demo and Savin-Williams (1981) suggested that the

unexplained variance indicated incongruence in measuring a unitary concept. Furthermore, as self-esteem measures were administered over a one week period, one might question whether or not aspects of self-esteem changed, thus influencing the magnitude of the convergent correlations. However, correlational data alone made it difficult to establish reasons for unexplained variance.

Delugach, Bracken, Bracken & Schicke (1992) assessed the convergent validity of three self-concept inventories (Self-Description Questionnaire I (SDQ I), Self-Description Questionnaire II (SDQ II) and the Multidimensional Self-Concept Scale (MSCS)). The sample included 43 fifth and sixth grade students. Results indicated moderately strong total score correlations between SDQ I and MSCS (r = .69) and SDQ II and MSCS (R= .80). Subscale score correlations ranged between r = .29 through .82 for SDQ I and MSCS and r = .36 through .74 for SDQ II and MSCS. Results indicated that all three measures were measuring the same construct.

Multidimensional self-concept validity was assessed by Bracken & Howell (1991) using Coopersmith's Self-Esteem Inventory (CSEI), Multidimensional Self-Concept Scale and the Piers-Harris Children's Self-Concept Scale (PHCSCS). A total sample of 65 fifth and sixth grade students completed all three self-concept/esteem measures. Result indicated that the all three measures were highly intercorrelated with correlations between CSEI and PHCSCS r = .83, CSEI and MSCS r = .73 and the MSCS and PHCSCS r = .85. Thus, Bracken & Howell (1991) concluded that the measures were assessing the same parameter.

Four studies, [Byrne, (1983), Hamilton, (1971), Silber & Tippett, (1965) and

Winnie, Marx & Taylor, (1977)], employed the multitrait-multimethod technique outlined by Campbell and Fiske (1959). Only the study by Hamilton (1971) was close to realizing the stringent criteria of Campbell and Fiske (1959). Hamilton (1971) used four methods each measuring self-esteem, dominance and dogmatism. These included self-report (California Psychological Inventory), self-rating (Janis-Field Feelings of Inadequacy Scale, a Self Rating Scale), behavioral checklist (Leary Interpersonal Checklist) and peer nomination. Subjects included 70 male undergraduate students. Hamilton (1971) noted high convergent correlations between self-report measures (r = .58 through .67), but only low to moderate correlations between self-report and peer-ratings of self-esteem (r =.23 through .33). In line with Savin-Williams and Jaquish (1981), the data suggested two conceptually distinct concepts were being measured. Hamilton's (1971) results bring further into question whether or not observation of the presented self is a viable method for measuring self-esteem. However, one must question whether a larger more heterogenous sample would have yielded similar results.

Silber and Tippett (1965) used three different self-report measures to validate self-esteem, (Rosenberg Self-Esteem Scale , Role Repertory Test [(Divided into measures of subjective satisfaction and the difference between self and ideal self)], Self Image Questionnaire). Forty four college students, 23 males and 21 females aged 17-21 from different colleges were included in the sample. Moderate to high correlations (r = .56 through .83) were obtained between all measures. However, one must question the use of all self-report (pencil and paper) measures, as common variance due to style of testing might account for such results.

Winnie, Marx and Taylor (1977) analyzed three multidimensional self-report measures of self-esteem (Sears Self-Concept Inventory, Gordon How I See Myself Scale and Piers-Harris Children's Self-Concept Scale). Subjects included 61 males and 42 females from grades 3-6. Strong evidence was obtained for convergent validity between subscales measuring aspects such as physical appearance, social, intellectual status and emotional self-perceptions. However, there was little evidence for discriminant validity due to the high multitrait-monomethod correlations which the authors suggest were due to shared method variance. Thus the use of similar methods of measurement within a multitrait-multimethod matrix provided limited discriminant validity, as method variance may have confounded the results.

Shavelson and Bolus (1982) examined the multifaceted hierarchical nature of self-concept using confirmatory factor analysis on self-report instruments measuring general self-concept (Piers-Harris Children's Self-Concept Scale, The Tennessee Self-Concept Scale), general academic, English, math, and science self-concept (The Michigan State Self Concept of Ability Scale), and school grades in English, math and science. Sixty nine males and 61 females from grades 7-8 completed all three measures. Shavelson and Bolus (1982) noted convergence on the two measures of general self-concept and on subscales over time (r. = .71 through .82 for time 1, and .70 through .83 time 2). Discriminant validity was provided by comparing the correlation between the two measures of general self-concept with the correlation between one general self-concept measure and school grades in English. Results indicated that discriminant validity was met with convergent validities being greater (r = .73 through .82) than the corresponding correlations in the

rows and columns. Evidence for the hierarchical nature of self-concept was obtained by comparing correlations between general self-concept with academic self-concept then specific school subjects self-concept, and finally with subject grades. In sum, general self-concept correlated highest with general academic self-concept, followed by school subjects then grades. As Shavelson and Bolus (1982: 16) suggested: " Self-concept is a hierarchical construct with general self-concept at the apex and situation specific concepts (at least as low in the hierarchy as subject matter specific self-concepts...) at the base."

Multi-trait multi-method analysis of the Coopersmith Self-Esteem Inventory (CSEI) general self and academic subscale, the Rosenberg Self-Esteem Scale and the Brookover Self-Concept Scale was provided by Byrne (1983). The criterion for convergent validity was met with theoretically linked measures providing moderate correlations (r = .42 through .60). Discriminant validity was met for most of the measures, with the CSEI academic subscale being the only exception. Overall the results suggested that the convergence of the self-concept measures provided evidence for their construct validity.

Flemming and Courtney (1984) provided factorial evidence for the dimensionality of self-esteem. Their sample consisted of 259 first year undergraduate students with a median age of 19. Factor analysis of the Janis-Field Inadequacy Scale (Self-Rating Scale) revealed five subscales; self-regard, social confidence, school abilities, physical appearance and physical abilities. Subscales were correlated to the Rosenberg Self-Esteem Scale, where convergent evidence was obtained (self-regard r = .78; social confidence r = .51; school abilities r = .35; physical appearance; r = .42; physical abilities r = .35; and a total

score for subscales r = .66). Fleming and Courtney (1984), suggested that the results indicate that the two measures were evaluating the same parameter, namely emotional self-esteem.

Using confirmatory factor analysis Demo (1985) analyzed four self-rating instruments of self-esteem (Rosenberg Self-Esteem Scale, Coopersmith Self-Esteem Inventory, Beeper Self-Reports, and Interviewee self-descriptions), peer ratings, and observer checklists (Williams and Jaquish (1981), The California Q-set (Form III)). Data were collected longitudinally throughout grades 9 and 10 form 24 males and 31 females. The two traditional questionnaires (Rosenberg Self-Esteem Scale and Coopersmith Self-Esteem Inventory) along with personal interviews provided convergent evidence for measuring the construct of experienced self-esteem, whereas ratings by external sources indicated measurement of presented self-esteem. No evidence was forthcoming as to the convergence of these two methods of measuring self-esteem, suggesting two distinct concepts.

Van-Tuinen and Ramanaiah (1979) utilized both a multitrait-multimethod and factor analytic design to investigate multiple self-report measures of global self-esteem (Tennessee Self-Concept Scale, Coopersmith Self-Esteem Inventory), social self-esteem (Janis-Field Feelings of Inadequacy Scale, Jackson Personality Inventory) and orderliness measures (Personality Research Form, Comrey Personality Scales). The samples were comprised of 97 male and 107 female undergraduate students. Results showed moderate to high correlations (.56 through .75) between measures of global self-esteem, indicating convergent validity. The criteria for discriminant validity was also met with measures of

orderliness not correlating highly (.07 through .34) with measure of global and social self-esteem. Factor analysis indicated that there were three factors, global self-esteem, social self-esteem and orderliness. The factor solution further enhanced the evidence for the multidimensionality of self-esteem with social self-esteem being a subconstruct of global self-esteem.

Marsh and Smith (1982), employed factor analysis to analyze the content of two self-report self-esteem inventories (The Sears Self-Concept Inventory and Coopersmith Self-Esteem Inventory). Data was also obtained on the stability of these measures over time. Subjects, (n = 91, interval = 1 year) were aged 8-13. Results indicated evidence of moderate convergent validity between the two scales (r = .42). However, little evidence was obtained for the discriminant validity of the subscales of these measures. Factor analysis of the Sears Self-Concept Scale revealed seven factors: physical ability, physical attractiveness, social relations (with same sex), work habits, school subjects, convergent mental ability and divergent mental ability. Factor analysis failed to provide any factor pattern for the Coopersmith Self-Esteem Scale. Furthermore, over a period of time the subscale scores for the Coopersmith Self-Esteem Inventory (mean r = . 39, for factor analytically derived scores) were less stable than the Sears Self-Concept Inventory (r = .57). In sum it appears that the inventories were not measuring the same construct and that the Coopersmith Self-Esteem Inventory had less internal consistency and stability than the Sears Self-Concept Inventory.

Multi-trait multi-method and confirmatory factor analyses was conducted on three multidimensional self-concept instruments (Self Description Questionnaire (SDQ), Piers-

Harris Children's Self-Concept Scale(PHCSCS) and the Perceived Competence Scale for Children (PCS)) by Marsh (1990 b). A total sample of 290 students with a mean age of 10.5 was used. Results from the multi-trait multi-method analysis provided support for the convergence of the three measures (mean r = .61). Evidence was also noted for discriminant validity with row and column correlations (mean r = .41) being lower than convergent validities. Confirmatory factor analysis noted lower convergent validities between the PCS and PHCSCS general physical scales. Moreover, the confirmatory factor analytic approach enabled trait factors to more readily identified. For example Marsh (1990) identified two physical factors on the PHCSCS, instead of the one originally reported by Piers (1984), that were related to the physical appearance and physical ability scale from the SDQ I. However only one physical factor, physical ability, was identified for the PSC. Thus Marsh (1990) concluded that separate traits for physical appearance and physical ability appeared to fit the data better than one general physical factor as used by the PCS and the PHCSCS.

Traits of general, academic, English and mathematics self-concept from four self-rating instruments based on Lickert scaling (The Self-Description Questionnaire III), semantic differential scale (The Affective Perception Inventory), Guttman scale of a global measure (Rosenberg Self-Esteem Scale) and academic facets of self-concept (Forms A,B,C of the Self-Concept of Ability Scale), were obtained by Byrne (1989). The sample included 252 low and 588 high ability children from grades 11-12. Data were subjected to multitrait-multimethod and confirmatory factor analyses. Convergent validity between all types of scales (r = .39 through .84) for both groups was evident. Method bias was

present for both groups. Furthermore, discriminability of traits was inconsistent across analyses. Whereas the multitrait-multimethod approach demonstrated convergence of measures, confirmatory factor analysis confirmed this for the high ability group only. Confirmatory factor analysis further indicated the high ability group differed on self-concept measured by Lickert and Guttman scales and that trait factors were less clear for the low ability group. Byrne (1989) suggested that these problems may be due to a strong trait-method interaction effect. However, Byrne, (1989), suggested that the analytical design used rendered it problematic to clarify the trait-method interactions effects. Overall both analyses found strong evidence for convergence of measures but contradictory evidence for divergent validity. Although the multitrait-multimethod analysis noted evidence for discriminability between traits across both groups, confirmatory factor analysis provided evidence for high ability students only. Based on these results the superiority of confirmatory factor analysis over multitrait-multimethod procedure was evident. Byrne (1989) suggested that confirmatory factor analysis provided "...more detailed evidence of construct validity within groups, and (b) testing for the equivalency of construct validity across groups." (p. 503)

The above studies indicate that there is evidence for the validity of unidimensional self-esteem measures (Byrne, 1983; Demo, 1985; Demo and Savin-Williams, 1983; Hamilton, 1971; Savin-Williams and Jaquish, 1981; Silber and Tippett, 1965) and multidimensional measures (Shavelson and Bolus, 1982; Winnie et al, 1977). There are studies that have addressed cross theoretical model validity based on unidimensional and multidimensional (Bracken & Howell, 1991; Flemming & Courtney, 1984; Van Tuinin &

Raminaiah, 1979; Marsh & Smith, 1982) and multidimensional and hierarchical (Delugach et al, 1992; Marsh, 1990) measures of self-esteem. However, no known studies have attempted to analyze whether self-esteem instruments derived from the different theoretical models provide convergent validity.

Conclusion

For James, (1890), Cooley, (1902), and Mead, (1934), self concept was a conscious introspective view of self mediated by the experiences of the individual. Self-concept for all three writers involved an evaluative component that accounted for differences in self-esteem. However, it was Rosenberg (1979), and Coopersmith (1967), who empirically accounted for self-esteem variation. Rosenberg (1979), and Coopersmith's (1967) work presented a framework from which theoretical models of self-concept could be tested. This led to three theoretical models of self-concept being forwarded. As a consequence, a vast array of self-esteem measures have been developed under the various theoretical self-concept models. Ultimately this has created confusion as to which measure one should adopt. The lack of a 'gold standard' measure of self-esteem has been the major drawback in self esteem research, as results using the various measures provide conflicting data. This is highlighted in the self-esteem and pediatric obesity literature.

Self-esteem is subject to individual experience. Thus unidimensional measures of self-esteem are inadequate in addressing the contributions of many situation-specific experiences. As obesity is a multifaceted problem (Williams, 1986) it makes sense to identify potential areas that may negatively impact on this disorder. Multidimensional

hierarchical models of self-concept not only present opportunities to address the underlying subdomains of self-concept, they further acknowledge the effect of individual experience. Locating areas of low-self-esteem can also identify sources in the negative experiences and direct the focus of the necessary interventions.

Convenient obesity indicators have also contributed toward the lack of consensus in the childhood obesity and self-esteem literature. The use of different obesity indicators has contributed to a deficiency in understanding the relationship between pediatric obesity and self-esteem. Thus, until more accurate field measures are made available and a cut-off point at which health risk can be defined is identified, data will continue to be confusing.

Validating self-esteem measures is problematic due to the limited methods one can adopt. As pencil and paper self-report methods are most commonly used, establishing construct validity as set out by Campbell and Fiske (1959) is difficult. Studies that have attempted to validate behavioral observation of self-esteem with self-report measures of self-esteem (Demo, 1985; Hamilton, 1971; Savin-Williams and Jaquish, 1981) found no congruence in results between methods. Thus two distinct concepts are evident (experienced and presented self-esteem), further highlighting the difficulty in establishing alternative methods of estimating self-esteem.

Ultimately there is evidence for construct validity between unidimensional measures of self-esteem (Byrne,1980; Demo, 1985; Demo and Savin-Williams, 1983; Hamilton, 1971; Savin-Williams and Jaquish, 1981; Silber and Tippett, 1965) and between multidimensional measures of self-esteem (Shavelson and Bolus, 1982; Winnie et al, 1977). However, cross theoretical model analyses of self-esteem measures provided

contradictory evidence, with some demonstrating convergence between theoretical models (Bracken & Howell, 1991; Bracken et al, 1992; Fleming and Courtney, 1984; Marsh, 1990; Van Tuinen and Raminaiah, 1979) and others suggesting no convergence (Marsh and Smith, 1982).

The lack of consensus in the pediatric obesity and self-esteem literature is partially due to measurement problems associated with theoretical foundations of self-esteem. Therefore caution is warranted as to the validity of the self-esteem measures used given the lack of cross theoretical model validation. Furthermore, as obese children are often compared to leaner children one must question whether self-esteem is equivalent across these two groups in light of Byrne's (1989) data. Until there is evidence of a 'gold standard' for the measurement of self-esteem and classification of obesity, the relationship between the two variables will continue to be contradictory. Thus there is a need to establish the validity of self-esteem measures based on the various theoretical models of self-concept, to provide a clearer picture as to whether obese children have low/high self-esteem. If self-esteem is a mediating variable in the onset, maintenance or reduction in obesity then establishing valid estimates of this construct is of paramount importance.

CHAPTER 3

Methods and Procedures

Subjects

All subjects were recruited from eight Edmonton Public Schools on the basis of obtained permission and convenience. Subjects included grades 8 and 9 obese and nonobese male and female students. Informed consent (See appendix A) was obtained from each subject and their parents/guardians. Ethics approval was obtained, prior to data collection, from the Ethics Committee, Faculty of Physical Education and Recreation. Permission was obtained from the Edmonton Public School Board and school principals prior to subject recruitment and testing. All potential subjects were informed as to the measures and procedures to be used before consent forms were distributed. Those students who wished to be considered were given consent forms to have completed and returned to the school principal's office.

Measures

Anthropometric measures included height, weight and a sum of five skinfolds (CSTF, 1986). Self-esteem measures used were RSES (Rosenberg, 1979), PHCSCS, (Piers, 1984) and the SDQ (Marsh, 1990a). All measures were appropriate for use with adolescent subjects.

Anthropometric Measures

Anthropometric data were collected by a research assistant of the same gender as the subject in the following manner. Height was taken with shoes removed to the nearest .2 cm using a set square and measurement tape. Weight was measured with subject in

indoor clothing on a beam balanced scale to the nearest .5 kilogram, with scales calibrated daily. Skinfolds (triceps, biceps, subscapular, iliac crest, medial calf) measures were obtained according to the procedures outlined in the CSTF (1986) manual. All anthropometric measures were collected by the same 2 trained research assistants throughout data collection period. An inter-individual reliability coefficient (r_{xx} = .97) for research assistants was established prior to the study by administering skinfolds on a sample of 30 subjects.

Classification of Obesity

Subjects whose sum of 5 skinfolds were equal to or greater than the 85th percentile of Canadian population norms (Canada Fitness Survey, 1985), were operationally defined as obese.

Self-Esteem Measurement

Subjects completed the RSES (Rosenberg, 1979), PHCSCS (Piers, 1984) and the SDQ (Marsh, 1990a). The three questionnaires took a total of approximately 45 minutes to complete. All items were read aloud to the subjects in intact classes by the same research assistant throughout the data collection. Random sequencing of the questionnaires from class to class was used to avoid systematic effects of one form on the other. A subset (n = 24) of the total sample was initially tested on all self-esteem measures during instrument administration and then retested 3 weeks later on the same three self-esteem measures to provide an estimate of the stability. For all self-esteem questionnaires, responses to negatively worded items were reversed so that the highest response code was indicative higher self-esteem.

Self-Esteem Measures

Rosenberg Self-Esteem Scale (Rosenberg, 1979). The RSES is an instrument based on a unidimensional model of self-esteem. The scale is comprised of 10 items utilizing a 4 point scale (strongly agree, agree, disagree, strongly disagree). Scores range from 10-40 with higher scores indicating higher self-esteem. Internal consistency of α = .88 has been reported by Flemming and Courtney (1984). Test-retest reliability coefficients between r_{XX} = .82 and .85 for 1 week and 2 weeks respectively have been obtained (Flemming & Courtney 1984; Silber & Tippett 1965)

Piers-Harris Children's Self-Concept Scale, (Piers, 1984). The PHCSCS is a multidimensional scale that measures child and adolescent self-concept. The scale is constructed of 80 statements with " Yes" or "No" choices to each item. Half the items are indicative of low self-esteem and the other half high self-esteem. High scores indicate greater self-esteem. Subscales are: anxiety (14 items), behaviour (16 items), intellectual and school status (17 items), physical appearance and attributes (13 items), popularity (12 items), and happiness and satisfaction (10 items). Internal consistency on total and subscale scores ranged between α = .72 through .90 (Piers, 1984). Test-retest reliability of between r_{XX} = 0.72 to 0.96 have been reported using this scale over two weeks and four months respectively (Blascovich & Tomaka, 1991).

Self-Description Questionnaire II (Marsh, 1990a). The SDQ is a measure of self-esteem based on heirarchical multidimensionality. The SDQ is constructed of 102 items that measure 11 subscales including: general self (10 items), mathematics (10 items), verbal (10 items), general school (10 items), physical abilities (8 items), physical

appearance (8 items), relations with same sex peers (8 items), relations with opposite sex (10 items), relations with parents (10 items), honesty (10 items), and emotional stability (10 items). Subjects respond on a 6 point Lickert scale (true, mostly true, sometimes true, sometimes false, mostly false and false). Higher self esteem is indicated by higher scores. Internal consistency for the subscales ranged between α = .87 through .94, and for test-retest coefficient r_{xx} = 0.72 to 0.88 over a 7 week interval (Marsh, 1990a).

Statistical Analyses.

Data were analyzed using the Statistical Package for the Social Sciences (SPSS)

Reliability coefficients for all three self-esteem measures were calculated using Cronbach's Alpha for internal consistency and Pearson's Product Moment Correlation Coefficient for test-retest stability.

To determine if the measures classified adolescent self-esteem differently, total scores for each subject were calculated for each of the three self esteem scales. Subjects were then categorized into high (top third), medium (middle third) or low (bottom third) self-esteem for each indicator. Data were analyzed using a chi-square log-linear analysis with backwards elimination.

To determine where gender differences occured on total scores the RSES, PHCSCS and SDQ were subjected to MANOVA. Separate MANOVA's were used to determine where gender difference occured on subscale scores on the PHCSCS and SDQ.

Obese subjects were age, gender, and school matched to form a subset of 126 (63 obese, 63 nonobese) subjects for secondary analyses. Differences between obese and nonobese boys and girls on RSES, PHCSCS and SDQ total mean scores were analyzed

using a MANOVA. Separate MANOVA's for the PHCSCS and the SDQ were performed to ascertain where subscale scores differed amongst obese and nonobese boys and girls.

To determine where significant interactions and main effects occurred, F- tests were calculated using individual mean square error terms (MS/ Error).

The $\alpha = .05$ level of significance was employed throughout the data analyses.

CHAPTER 4

Results and Discussion

Subject Characteristics

Of 1,835 parental consent forms distributed a total 349 (19 %) students returned forms. Due to missing data a final pool of 322 (92.2%) subjects was obtained for all analyses (Table 2). A total of 63 (19.6%) students (36 females, 27 males) were classified as obese according to sum of five skinfolds (CSTF, 1986) (Table 3). A list was generated for all obese and potential matched control subjects. Obese subjects were randomly age, gender, and school matched with a control from the list to form a subgroup of n = 126 (Table 3).

Table 2 Group Descriptives for Gender

	N	Mean	SD
Female	180		
Age		14.36	.62
Height		161.64	6.28
Weight		54.35	10.50
SOS		69.09	27.26
Male	142		
Age		14.55	.68
Height		168.78	8.63
Weight		58.22	11.08
SOS		48.23	21.78
Total Group	322		
Age		14.44	.65

SOS = Sum of five skinfolds

No statistically significant differences for height were observed between obese and nonobese subjects. However, obese compared to nonobese subjects were significantly heavier (t = 8.27: P = .000), and fatter (t = 13.82: P = .000). No statistically significant

differences were observed for height between obese and nonobese boys but obese boys were statistically significantly heavier (t = 4.29: P = .000) and fatter (t = 10.89: P = .000) than their nonobese peers. Similar results were observed for obese girls compared to nonobese girls with no statistically significant difference in height, but obese girls were heavier (t = 7.40: P = .000) and fatter (t = 11.75: P = .000) than nonobese girls.

Table 3 Descriptives for Matched Obese and Nonobese Subjects

	N	Mean	SD
Female			
Obese	36		
Height		161.55	5.10
Weight		67.48	12.83
SOS		111.44	29.25
Nonobese	36		
Height		161.77	6.57
Weight		49.93	5.75
SOS		50.75	10.25
Total Females	72		
Height		161.64	7.28
Weight		58.60	13.18
SOS		81.10	37.51
Male	27		
Obese			
Height		169.63	7.07
Weight		67.67	11.46
SOS		84.10	24.10
Nonobese	27		
Height		169.99	9.41
Weight		55.49	9.32
SOS		33.18	3.10
Total Males	54		
Height		168.78	8.63
Weight		61.58	12.03
SOS		58.64	30.82

Reliability

Reliability does not necessarily imply validity, but it is a condition that must be met before validity can be examined and claimed (Pedhazur & Schmelkin, 1991). All measures were subjected to two reliability procedures: internal consistency and test-retest. Internal consistency refers to the ability of a measure to provide consistent patterns of relationships between items purporting to measure the construct under investigation (Traub, 1994). Cronbach's alpha is considered a measure of such consistency (Traub, 1994), and was calculated on all total and subscale scores where appropriate. Wylie (1989) noted that reliability estimates should be in excess of .90 to be acceptable as a measure of good internal consistency. However, as Pedhazur (1991) suggested, reliability estimates must be viewed in conjunction with test use. That is, if a test is to be used for making important decisions about an individual then reliabilities need to be high. Test-retest reliability refers to the ability of a measure to provide consistent results over a period of time, therefore stability is the key element. An important consideration for test-retest reliability coefficients is the time lapse between testing. Shorter periods of time between tests warrant caution as high reliability values may result from extraneous sources of variance such as memory effects. To assess test-retest reliability Pearson Product Moment Correlation Coefficients were calculated on the same subsample using the same instruments twice three weeks apart.

Table 4 displays the internal consistency coefficients for the RSES total score, SDQ and PHCSCS subscales and total scores for the total sample. Total scores for the PHCSCS (α = .92) and the SDQ (α = .96) indicated strong internal consistency and

concurred with the values of α = .90 reported by Piers (1984) and α = .94 by Marsh (1990a). However, the RSES yielded a moderately high value (α = .64) which was lower

Table 4 Internal Consistency and Standard Errors of Measurement for RSES, PHCSCS and SDQ

	α	SE_M
RSES Total Score	.64	2.31
PHCSCS		
Anxiety	.85	1.16
Behaviour	.83	1.17
Intell/School	.80	1.70
Phy App	.83	0.95
Popularity	.78	0.96
Happy	.84	1.07
Total PHCSCS Score	.92	4.78
SDQ		
Math	.94	3.34
Phy App	.92	2.65
Gen Self	.92	2.66
Honesty	.87	3.19
Phys Ability	.85	3.18
Verbal	.87	3.72
Emotional	.87	3.70
Parents	.89	2.87
Gen School	.92	2.95
Same Sex Rel (M)	.88	2.66
Opp Sex Rel (M)	.94	2.43
Same Sex Rel (F)	.91	2.80
Opp Sex Rel (F)	.93	2.58
Total SDQ Score	.96	13.03

Note: N = 322. SE_M Based on the formula $SE_M = \sqrt{1 - r_{xx}}$.
Subscale scores for Same sex and opposite sex relationships are presented separately for males (n = 142) and females (n = 180) for SDQ.

than those noted in Wylie's (1989) review of the literature who reported values ranging from α = .74 through .87. The internal consistencies for the PHCSCS subscales were

high, ranging from α = .78 through .85. The SDQ subscales were slighly more internally consistent, with values ranging from α = .85 through .94 and were in agreement with Piers' (1984) own results. These results are, however, lower than the minimum acceptable value of .90 for internal consistency suggested by Wylie (1989).

The PHCSCS yielded a standard error of measurement of 4.78 for the total score which was similar to that reported by Piers (1984), whereas standard errors for the subscales (SE_M = .95 through 1.61) were lower than those reported by Piers (1984). The standard error for the SDQ (SE_M = 13.03) total score was lower than the value given by Marsh (1990a) as were the subscale standard errors of measurement. The RSES standard error (SE_M = 2.1) could not be compared to the previous literature as published results could not be found.

Test-retest coefficients for a subsample of n = 24, tested twice, three weeks apart for RSES total score and the PHCSCS and SDQ total and subscale scores are shown in Table 5. The total score test-retest coefficient for RSES (r_{12} = .74) was within those reported by Silber and Tippett (1965) (r_{12} = .85), and Byrne (1983) (r_{12} = .65). Stability coefficients for PHCSCS total score (r_{12} = .86) were higher than the median test-retest score of .73 given by Piers (1984) in the test manual. As was stated in Wylie's (1989) review, no subscale test-retest reliabilities for the PHCSCS were found for comparison. SDQ test-retest total and subscale results were also comparable with those reported by Marsh (1990a). Overall the test-retest values obtained with this sample using the PHCSCS and the SDQ were high, indicating strong stability. However, the lower test-

retest coefficient obtained on the RSES suggested weaker stability over time.

Table 5 Internal Consistency and Test-Retest Total Score for RSES and Total and Subscale Scores for PHCSCS and SDQ

	r_1	r_2	r_{12}
RSES Total	.61	.62	.74
PHCSCS			
Anxiety	.71	.78	.80
Behaviour	.81	.78	.88
Intell/school	.79	.91	.77
Phys App	.88	.85	.92
Popularity	.77	.85	.87
Happy	.84	.81	.81
PHCSCS Total Score	.92	.94	.86
SDQ			
Math	.96	.96	.80
Phy App	.93	.94	.90
Gen Self	.93	.94	.88
Honesty	.84	.68	.81
Phy Ability	.78	.74	.87
Verbal	.85	.88	.93
Emotional	.89	.90	.90
Parents	.89	.94	.89
Gen School	.93	.91	.88
Same Sex Rel (M)	.88	.95	.62
Opp Sex Rel (M)	.88	.93	.87
Same Sex Rel (F)	.89	.89	.91
Opp Sex Rel (F)	.94	.95	.92
SDQ Total Score	.97	.96	.92

Note: N = 24 (Males = 8; Females = 16). r_1 = alpha coefficient for test scores at time 1: r_2 = alpha coefficient for test scores at time 2: r_{12} = test- retest reliability. Subscale scores for Same sex and opposite sex relationships are presented separately for males and females for SDQ

Discussion of Reliability

Overall the PHCSCS and the SDQ demonstrated good internal consistency. Although the RSES yielded lower alpha coefficients the result is, considering the number of items, acceptable. The lower value may also be a function of interpretive answering of

the items in the RSES. As the items are global and less specific than the PHCSCS and SDQ, looser interpretation of the questions may have affected its internal consistency. The PHCSCS and the SDQ on the other hand provided high internal consistency, with the SDQ demonstrating slightly higher alpha coefficients. However, this is to be expected given that the reliability of a test is partially affected by the of number of items in the instrument. Furthermore, the personal and specific nature of the questions lead to more consistent patterns of responding. Subscales permit individuals to draw on the previously associated questions allowing a more dependable answering pattern. Standard errors of measurement for the PHCSCS and SDQ subscales were generally lower than the values reported by Piers (1984) and Marsh (1990a). Standard error of measurement is a partial function of the internal consistency. Higher alpha coefficients are associated with lower standard errors and are therefore more indicative of reliable measures (Traub, 1994). Given the moderately high to high alpha coefficients for total and subscale scores for all three self-esteem measures, the relatively low standard errors were to be expected .

Test-retest coefficients indicated that for total scores the SDQ and the PHCSCS were more stable over time than the RSES. The PHCSCS and SDQ subscale test-retest coefficients were again strong and were higher than the values reported by Piers (1984) and Marsh (1990a) respectively. Overall these data provide strong support for the reliability of the PHCSCS and the SDQ, but indicate only moderately strong reliability for the RSES.

Validity

Loglinear Analysis

Loglinear analysis is a statistical technique that is employed in the analysis of categorical variables. As Kennedy (1983) stated, "Loglinear analysis may be viewed as the issue of a traditional marriage between the ANOVA and chi-square goodness of fit." Thus the ANOVA approach of model fitting is applied to the goodness of fit of the expected cell frequencies to the observed cell frequencies. One selects the model that provides the most acceptable fit to the observed data. If the model selected provides an adequate fit to the data then the chi-square will be low. Conversely if the model does not provide any adequate fit to the data, higher chi-square will be observed. With backwards elimination of the threeway and twoway interactions effects from the full model, increases in chi-square are associated with increases in the significance level. In other words the effects of what is left out of the model significantly contribute towards whether the model does not fit the data. Thus removing effects from the model allows one to assess how well the information left in the model fits the data.

Table 6 shows the best model to fit the data for the chi-square loglinear analysis with backwards elimination. Results indicated that the three self-esteem measures do not classify subjects in a similar pattern. No third or second order effects when removed from the model provided an adequate fit of the model to the data. The RSES appears to be the measure that was contributing most to the lack of model fit, as when the SDQ x PHCSCS interaction effect was removed from the model, the greatest increase in chi-square (χ^2 = 171.57; P = .000) was observed indicating that the SDQ and PHCSCS were more highly

related than the RSES and PHCSCS or RSES and SDQ. Examination of Model (3.1) and Model (3.2) suggested that when the SDQ x PHCSCS effect is left in the model then the increase in chi-square was less than Model (3.3). Thus although the three measures provided different patterns of relationships, the instrument that contributed considerably to the poor model fit was the RSES.

Table 6 Chi-Square Loglinear Analysis with Backwards Elimination for RSES, SDQ and PHCSCS.

Full Model (1)			$(\chi^2 = 0)$	$(df = 0)$
	Drop	*Left*	Δ inχ^2	Δ in *df*
Model (2)	RSP	RS RP SP	15.44*	8
Model (3.1)	RP	RS SP	38.54**	4
Model (3.2)	RS	RP SP	15.68**	4
Model (3.3)	SP	RS RP	171.57**	4

Note N = 322. (* < .05)(** < .01). R = RSES, S = SDQ and P = PHCSCS.

This finding is confirmed by examining the percentage agreement between each measure. As noted in tables 7, 8, and 9, the percentage agreement between the SDQ and PHCSCS (71.4%) was higher than that found between PHCSCS and RSES (56.8%) and SDQ and RSES (53.7%).

Table 7 Cross Tabulation of Agreement Between SDQ and PHCSCS

		SDQ			Total
		Low	Med	High	
PH	Low	85	21	1	107
	Med	20	61	23	104
	High	2	25	84	111
Total		107	107	108	322

Table 8 Cross Tabulation of Agreement Between PHCSCS and RSES

		RSES			Total
		Low	Med	High	
PH	Low	74	20	13	107
	Med	20	40	44	104
	High	8	34	69	111
Total		102	94	126	322

Table 9 Cross Tabulation of Agreement Between SDQ and RSES

		RSES			Total
		Low	Med	High	
SDQ	Low	63	28	16	107
	Med	33	37	37	107
	High	6	29	73	108
Total		102	94	126	322

Table 10 shows the Spearman rank order correlations between RSES, PHCSCS and the SDQ. The results indicate moderately strong association between the RSES and PHCSCS (r = .63) and SDQ (r = .62). However, the correlation between PHCSCS and SDQ was high (r = .85).

Table 10 Spearman Rank Order Correlations Between RSES, PHCSCS and SDQ

	RSES	PHCSCS	SDQ
RSES	1.000		
PHCSCS	.6396**	1.000	
SDQ	.6271**	.8528**	1.000

Note: n = 322, **p < .000

Discussion of Validity

Overall the results suggest that the three measures do not classify individuals in a consistent pattern. This is a concern since these measures purport to assess the same construct, namely self-esteem. No one theory of self-concept/esteem has been dominant in the literature and there is a lack of consensus as to which best represents self-esteem. The plethora of self-esteem measures available is testament to this. These results indicate that classifaction of subjects, based on the various theoretical models of self-esteem to be inconsistent. However, the RSES appeared to be the instrument that was contributing most towards the poor congruency of relationships between all three measures. The correlations between the RSES and the PHCSCS and SDQ support the loglinear analysis that the RSES was influencing the poor patterns of relationship among the three instruments. This is not to say that the PHCSCS and the SDQ provided any better

patterns of relationship, but were more in agreement with one another. This was further highlighted by the high correlation between the PHCSCS and the SDQ indicating a high degree of association. In other words, although the PHCSCS and the SDQ were not able to provide any consistent patterns of relationship, there is some evidence that they may be measuring a similar underlying construct (or at least classifying it in the same way).

Some potential reasons for differences in classification between three measures of self-esteem are presented below.

Lower internal reliability of a test will contribute to its lack of convergence with the other instruments purporting to measure the same construct (Bracken, 1988). As a result standard errors of measurement will be larger and as Bracken (1988) stated..."produce large confidence intervals surrounding the examinee's 'true score'." (p.161) As the RSES had the lowest internal reliability of all the measures,then this may account for some of the discrepancy in its ability to classify subjects in a similar pattern to the other 2 measures.

Content differences on self-esteem inventories also impact on the performance of tests claiming to measure the same construct (Bracken, 1988). One must question the content of the RSES. The items lack specificity. The global nature of the items, such as "I feel that I have a number of good qualities", requires more information to be integrated into the answering process. Not only does the individual have to perceive the things that they feel good about, they also have to provide some internal summation of these in order to answer the question. The interpretive and global nature of the items may have compromised the performance of the RSES, as some of the subjects did ask the author to

clarify what certain questions from the RSES meant during the tests administration. In sum, the lack of specificity and the amount of information processing required may result in an inaccurate representation of self-esteem. Global questions therefore may not be representative of an individual's level of self-esteem as certain key factors may not be integrated into the response. Thus there is a lack of focus on more specific dimensions.

Of the two multidimensional instruments the SDQ provided the highest internal and test-retest reliability coefficients. This may be due to its sound theoretical base and clear factor structure (Marsh, 1990a). The lower reliability of the PHCSCS in comparison to the SDQ may be due to its weaker factor structure. The PHCSCS was not originally based on the Shavelson et al (1976) model but subsequent revisions have indicated six subscales, although these are not clearly defined (Piers, 1984). Certain items in the PHCSCS are used in more than one subscale and some are only used in the total score, this may lead to ambiguity in answering questions and contribute to the weaker factor structure. Subscales are therefore not mutually exclusive on items. For example the item " I am smart" loads on both intellectual and school status, and physical appearance and attributes. How an individual perceives the word "smart" will influence their response. A child, for example, may perceive himself/herself as being well dressed but not very intelligent. Such interpretation of this and other items may ultimately affect the total and subscales scores for an individual and lead to a false impression of what the actual level of self-esteem is. This is a serious flaw when it comes to identifying those with low self-esteem as one is not sure of the meaning attached to some of the responses. The unclear factor structure of the PHCSCS is a weakness and does impact on its

performance when used in conjunction with other well defined factored measures such as the SDQ.

The clear factor structure of the SDQ indicated that the items within each subscale were more homogenous and required less interpretation than the PHCSCS and the RSES, as each only loads on one subscale. As a result the SDQ is able to reduce ambiguity in responses and provide a more consistent and reliable answering pattern (Marsh, 1990a).

Publication dates of the tests will also produce results that differ for measures that are designed to assess the same construct (Bracken, 1988). The RSES was originally published in 1965, the PHCSCS in 1964 and the SDQ II in 1990, thus 26 years span these measures. Normative data were found for the PHCSCS and the SDQ. Scores at the 33rd percentiles for the PHCSCS and the SDQ from this study were compared to the published normative data. In the case of the PHCSCS the 66th percentile score (69) was equivalent to the 91st percentile of the normative tables. Similarly, for the SDQ the 66th percentile score (493) was equivalent to 78th percentile of the normative tables. Therefore, 33% of this sample was classifed as having high self-esteem, as compared to only 9% (PHCSCS) and 22% (SDQ). Thus, more children in the present study (33% compared to 9% [PHCSCS] or 22% [SDQ]) were classified as having high self-esteem when the raw score for each measure was compared to the percentile tables. This may be due to self-esteem promotion and enhancement in schools and also the self-selection of subjects into the study. Given the poor response rate in returning consent forms, then those subjects who took part in the study, may have been those who were interested in self-esteem. However, one must consider that the norms, especially those for the PHCSCS which are over 26

years old, may not be representitive of todays population. Therefore the changing characteristics of the population over time may account for the differences in these results. Furthermore, as the SDQ II was developed in Sydney, Australia, the PHCSCS in Pennsylvania, U. S.A., and the RSES in New York State, U.S.A., then cultural variation when developing the test may account for test score differences. Using measures developed in different geographic regions with Canadian adolescents may account for a portion of the incongruent pattern of relationships obtained using these instruments.

MANOVA

Separate MANOVA's were calculated to determine differences between boys and girls on total and subscales scores for all three self-esteem measures. MANOVA's were also used to determine the differences between obese and nonobese boys and girls on total and subscale scores for all three self-esteem measures.

Table 11 contains the mean RSES total score and PHCSCS and SDQ total and subscale scores for the entire sample separated by gender. MANOVA results indicated a significant gender effect for RSES, PHCSCS and SDQ mean total scores (Pillais $F_{3,318}$ = 6.24; P = .000). F-tests indicated girls to be lower in mean total self-concept scores on RSES ($F_{1,320}$ = 15.8; P = .001), PHCSCS ($F_{1,320}$ = 5.52; P = .05) and SDQ ($F_{1,320}$ = 10.62; P = .01). MANOVA for the PHCSCS subscales also revealed significant gender differences (Pillais $F_{3,318}$ = 9.71; P = .000). Girls reported statistically significantly lower mean scores for anxiety ($F_{1,320}$ = 31.24; P = .001), happiness and satisfaction ($F_{1,320}$ = 10.26; P = .01) and physical ($F_{1,320}$ = 6.25; P = .05) subscales. No significant differences between boys and girls were observed for the PHCSCS behavior, intellectual and school

status, and popularity subscales.

MANOVA results indicated significant gender differences on the SDQ (Pillais $F_{11,310} = 6.92$; $P = .000$). Girls reported statistically significantly lower mean scores on emotion ($F_{1,320} = 24.56$; $P = .001$), general self ($F_{1,320} = 13.86$; $P = .001$), mathematics ($F_{1,320} = 4.84$; $P = .05$), opposite sex relationships ($F_{1,320} = 10.72$; $P = .01$), physical ability ($F_{1,320} = 18.58$; $P = .01$) and physical appearance ($F_{1,320} = 20.82$; $P = .001$) subscales. However, the mean score for the verbal (univariate $F_{1,320} = 6.94$; $P = .01$)

Table 11 Means and Standard Deviations for Total RSES Score and Total and Subscale Scores for the PHCSCS and SDQ for Boys and girls

	Boys (N = 142)		Girl (N = 180)	
	Mean	SD	Mean	SD
RSES Total Score	28.02	3.32	26.37**	4.09
PHCSCS				
Anxiety	10.79	2.83	8.61**	3.98
Behaviour	12.55	3.46	12.94	3.36
Intell/School	12.35	3.40	12.33	3.79
Physical	9.52	3.20	8.57*	3.51
Popularity	8.71	2.64	8.32	2.76
Happy	8.52	1.86	7.64**	2.83
Total Score	62.45	13.17	58.43*	16.72
SDQ				
Math	40.85	13.56	37.33*	14.82
Phy App	35.80	7.86	31.10**	10.08
Gen Self	51.94	7.50	48.03**	10.52
Honesty	46.49	9.48	47.08	8.63
Phys Ability	40.25	7.16	36.37**	8.64
Verbal	41.23*	11.11	44.34	10.37
Emotional	43.25	9.48	37.55**	10.92
Parents	38.23	7.64	36.77	9.34
Gen School	45.88	11.30	46.46	10.54
Same Sex Rel	52.04	7.98	50.41	9.40
Opp Sex Rel	36.41	9.95	32.70**	10.18
Total Score	472.42	63.03	448.60**	71.92

* P < 0.05, ** P < 0.01 are on F-tests.

subscale was statistically significantly higher for girls than boys. No significant differences for the honesty, parent relations, general school and same sex relations on the SDQ subscales were observed between boys and girls.

A subsample was gender, age, and school matched for obesity and MANOVA's were calculated to determine whether obese subjects differed from nonobese subjects on mean total and subscale self-esteem scores.

Table 12 contains the mean total and subscale scores between obese and nonobese subjects. The MANOVA main effect on total score for obesity status was non significant (Pillais $F_{3,120}$ = 2.46; P = .066). However, univariate F-tests indicated differences between obese and nonobese on RSES (univariate $F_{1,122}$ =5.88; P = .017) and PHCSCS (univariate $F_{1,122}$ = 5.74; P = .031) total scores. No differences between obese and nonobese were found with the SDQ total score.

The MANOVA results for the PHCSCS subscale scores indicated a significant main effect for obesity status (Pillais $F_{6,117}$ = 2.27; P = .041) between obese and nonobese, with mean subscale scores for happiness and satisfaction ($F_{1,122}$ = 6.14; P = .05), popularity ($F_{1,122}$ = 11.39; P = .01) and physical ($F_{1,122}$ = 8.61; P = .01) being lower in the obese group. No significant differences were noted for the anxiety, behaviour, and intellectual and school status PHCSCS subscales between obese and nonobese subjects.

For the SDQ a significant main effect for obesity status was observed on subscales (Pillais $F_{11,112}$ = 2.36; P = .012) on MANOVA. Obese subjects reported statistically significantly lower mean subscale scores on general self-concept ($F_{1,122}$ = 8.92; P = .01), physical appearance ($F_{1,122}$ = 8.24; P = .01), physical ability ($F_{1,122}$ = 12.73; P = .001) and

same sex relationships ($F_{1,122} = 4.55$; $P = .05$). No differences were noted on the remaining SDQ subscales between obese and nonobese subjects.

A comparison of nonobese boys and girls revealed no significant differences on total scores for the RSES, PHCSCS and the SDQ. No significant differences were noted between nonobese boys and girls on the PHCSCS and SDQ subscales.

MANOVA results for RSES, PHCSCS and SDQ total scores revealed no significant interaction between obesity status and sex (Pillais $F_{3,120} = 1.92$; $P = .129$) (Table 13). However, univariate F-tests indicated interactions on the RSES (univariate $F_{1,122} = 4.1$; $P = .045$) and SDQ (univariate $F_{1,122} = 4.6$; $P = .035$) total scores, but not for the PHCSCS. Given that girls exhibited lower self-esteem than boys on both total and certain subscale scores it was of interest to test whether obese girls were the group that were influencing the interactions on the univariate F-tests. Therefore F-tests were performed using individual MS/Error to ascertain where differences occurred. Obese girls reported statistically significantly lower mean total score than obese boys on the RSES ($F_{1,122} = 13.72$; $P = .001$) and SDQ ($F_{1,122} = 13.42$; $P = .001$) total scores. No significant differences were noted on the PHCSCS total mean score between obese boys and girls. Differences between obese boys and girls for mean PHCSCS and SDQ subscales scores are given in Table 13. For the PHCSCS, obese girls reported statistically significantly lower anxiety ($F_{1,122} = 19.41$; $P = .001$) and popularity ($F_{1,122} = 4.10$; $P = .05$) subscale scores than obese boys. No significant differences were noted on the remaining 4 PHCSCS subscales between obese boys and girls. Mean subscale scores for the SDQ revealed statistically significant differences with boys reporting higher scores for general

self-concept ($F_{1,122}$ = 21.31; P = .001) and physical ability self-concept ($F_{1,122}$ = 18.25; P = .001). Significant differences were not noted on the reminder of the SDQ subscales between obese boys and girls.

Table 12 Means and Standard Deviations for the RSES Total Score and PHCSCS and SDQ Total and Subscale Scores for Matched Obese and Nonobese Subjects

	Obese (N = 63)		Nonobese (N = 63)	
	Mean	**SD**	**Mean**	**SD**
RSES Total Score	25.44**	4.91	27.46	3.63
PHCSCS				
Anxiety	8.47	4.29	10.17	3.31
Behaviour	12.33	3.48	13.00	3.39
Intell/School	11.39	4.25	12.04	2.97
Physical	7.70**	3.84	9.12	3.17
Popularity	7.38**	3.17	9.00	2.20
Happy	7.20*	3.10	8.34	2.21
Total Score	54.85*	18.64	61.69	13.18
SDQ				
Math	38.03	14.46	36.39	15.70
Phy App	29.26**	11.18	34.00	8.32
Gen Self	46.36**	12.67	51.44	7.07
Honesty	47.38	8.56	46.14	10.21
Phys Ability	34.26**	10.19	39.47	6.92
Verbal	44.01	11.19	41.33	9.64
Emotional	38.39	12.30	42.01	9.07
Parents	35.57	9.18	38.61	8.76
Gen School	44.79	12.77	45.90	10.07
Same Sex Rel	47.98*	12.16	51.80	7.50
Opp Sex Rel	32.14	12.48	35.57	9.18
Total Score	438.22	90.09	462.71	60.13

* $P < 0.05$, ** $P < 0.01$ for F-tests.

No statistically significant differences between obese and nonobese boys were observed on mean RSES, PHCSCS and SDQ total scores. F-tests for obese and nonobese boys indicated no statistically significant differences on the PHCSCS and SDQ subscales.

A comparison of mean total and subscales scores and standard deviations on the RSES, PHCSCS and the SDQ for obese and nonobese girls are presented in Table 14. Obese girls reported significantly lower total scores on RSES ($F_{1,122} = 11.49$; $P = .01$) and SDQ ($F_{1,122} = 6.28$; $P = .05$) compared to nonobese girls. No significant differences were observed on the PHCSCS total score between obese and nonobese girls.

Table 13 Means and Standard Deviations for Total RSES Score and PHCSCS and SDQ Total and Subscale Scores for Matched Obese Girls and Obese Boys

	Obese Girls (N = 36)		Obese Boys (N = 27)	
	Mean	**SD**	**Mean**	**SD**
RSES Total Score	23.77**	4.29	27.66	4.00
PHCSCS				
Anxiety	7.02**	4.34	11.03	3.02
Behaviour	12.02	3.51	12.74	3.48
Intell/School	10.58	4.49	12.48	3.64
Phy App	6.97	3.80	8.88	3.67
Popularity	6.77*	3.23	8.18	2.49
Happy	6.25	3.40	8.48	2.08
PHCSCS Total Score	49.63	18.17	61.81	16.33
SDQ				
Math	33.77	14.43	43.71	12.63
Phy App	25.22	11.21	34.66	8.73
Gen Self	41.55**	13.41	52.77	8.10
Honesty	46.75	9.10	48.33	7.88
Phys Ability	30.44**	10.13	39.37	7.88
Verbal	43.13	11.78	45.18	10.47
Emotional	34.44	12.18	43.66	10.52
Parents	33.86	10.16	37.85	9.10
Gen School	42.88	13.69	47.33	11.17
Same Sex Rel	46.13	13.93	50.44	8.97
Opp Sex Rel	30.75	12.95	34.00	11.80
SDQ Total Score	408.97**	91.33	477.22	73.23

* P < 0.05, ** P < 0.01 on F-tests.

MANOVA results indicated no obesity status and sex interaction effect on the PHCSCS subscales (Pillais $F_{6,112} = 2.08$; $P = .060$) . However, univariate F-tests indicated

statistically significant interactions on the anxiety (univariate $F_{1,122} = 4.19$; $P = .043$) and popularity (univariate $F_{1.122} = 4.19$; $P = .043$) subscales. Therefore, F-tests were

Table 14 Means and Standard Deviations for Total RSES Score and PHCSCS and SDQ Total and Subscale Scores for Matched Obese Girls and Nonobese Girls

	Obese Girls (N = 36)		Nonobese Girls (N = 36)	
	Mean	**SD**	**Mean**	**SD**
RSES Total Score	23.77**	4.29	27.03	3.63
PHCSCS				
Anxiety	7.02**	4.34	9.58	3.85
Behaviour	12.02	3.50	13.00	3.50
Intell/School	6.25	3.40	8.02	2.64
Phy App	6.97	3.80	8.61	3.20
Popularity	6.77**	3.23	9.25	2.00
Happy	6.25	3.40	8.02	2.64
PHCSCS Total Score	49.63	18.17	60.66	15.14
SDQ				
Math	33.77	14.43	34.08	15.70
Phy App	25.22	11.21	32.58	8.75
Gen Self	41.55**	13.41	50.75	7.09
Honesty	46.75	9.10	46.75	8.62
Phys Ability	30.44**	10.13	39.19	6.78
Verbal	43.13	11.78	43.13	8.17
Emotional	34.44	12.18	41.11	11.13
Parents	33.86	10.16	37.88	9.98
Gen School	42.88	13.69	46.38	9.33
Same Sex Rel	46.13	13.93	51.30	8.02
Opp Sex Rel	30.75	12.95	34.52	9.16
SDQ Total Score	408.97*	91.33	457.72	60.83

* P < 0.05, ** P < 0.01 on F-tests.

performed using MS/Error terms, to determine where differences occurred. Obese girls reported statistically significantly lower mean anxiety ($F_{1.122} = 7.17$; $P = .01$) and popularity ($F_{1.122} = 15.01$; $P = .001$) scores than nonobese girls but not on the remaining PHCSCS subscales. The girls mean score difference scores were significantly greater than

the boys mean score differences for both anxiety ($F_{1,122} = 8.14$; $P = .001$) and popularity ($F_{1,122} = 18.10$; $P = .001$).

No significant interaction for obesity status and sex was noted on the SDQ subscales (Pillais $F_{11,112} = 1.59$; $P = .110$). Subscale univariate F-tests indicated differences on general self (univariate $F_{1,122} = 7.82$; $P = .006$) and physical ability (univariate $F_{1,122} = 7.85$; $P = .006$). Obese girls were observed as having statistically significantly lower mean subscale scores than their leaner peers on the general self ($F_{1,122} = 17.52$; $P = .001$) and physical ability ($F_{1,122} = 20.55$; $P = .001$) subscales. Mean score differences for girls were significantly greater than boys for both general self ($F_{1,122} = 49.88$; $P = .001$) and physical ability ($F_{1,122} = 31.00$; $P = .001$).

Discussion of MANOVA Results

Piers (1984) found no evidence of gender differences in global self-concept. This view was later supported by Wylie (1989) in her extensive review of literature. Hattie (1992;180) further concluded that.. "the majority of studies lead to the strong conclusion that sex of subject is not an important moderator when using self-concept scales." However, the weak gender effects in global self-esteem may have been the result of gender differences being counterbalanced due to the global nature and variety of the instruments used (Wylie 1979; Marsh 1990 a). Conversely, Marsh (1990a), observed significantly higher total, physical ability, physical appearance and math self-concept scores in males than females, but females demonstrated significantly higher scores on the verbal subscale. The data in the present study support Marsh's (1990a) observation of lower total self-esteem in girls than boys and for lower subscale scores in certain areas and

higher verbal subscale scores for females.

The effect for obesity status on RSES and PHCSCS total mean scores was statistically significant. Results are in contradiction to those reported by others (Kaplan & Wadden, 1986; Mendolson & White, 1982; O'Brien et al, 1990; Wadden et al, 1984) who observed no differences in self-esteem between obese and nonobese children.

The results from the various MANOVA's provide strong support for significant differences in mean subscale scores with nonobese subjects reporting some statistically significantly higher mean scores. The results for differences between obese and nonobese subjects on the PHCSCS and the SDQ subscales provided an in depth analysis of where low self-esteem was being exhibited. For the PHCSCS significant differences were observed with obese subjects reporting lower mean scores on the physical, popularity and happiness subscales. Obese compared to nonobese subjects reported lower mean scores on the physical appearance, physical ability, general self and same sex relations for the SDQ subscales. The lower subscale mean scores for obese subjects are in agreement with those who have reported lower self-esteem in certain areas in obese children (Corbin et al, 1995; Klesges et al, 1992; Manus & Killeen, 1995; Strauss et al, 1985).

No significant differences were noted on total self-esteem scores for the RSES, PHCSCS and the SDQ, or on the PHCSCS and SDQ subscales between nonobese boys and girls.

Obese girls reported significantly lower mean total self-esteem scores than obese boys on the RSES and SDQ. Lower mean subscale score for obese girls were reported in table 10 for the PHCSCS anxiety and popularity subscales compared to obese boys. Also

general self and physical ability mean subscale scores on the SDQ were significantly lower for obese girls. Results suggest that obese boys do not suffer from lower self-esteem. However, one must consider that obese boys may not be subjected to the same sociocultural pressures as obese girls such as 'ideal body size and shape' (Kilbourne, 1994). For obese boys size may in fact be advantagous in maintaining their self-esteem, given the high profile of sports where size is an important factor, such as linemen in gridiron football.

However, obese boys reported comparable mean total self-esteem scores with their nonobese counterparts on both the RSES, PHCSCS and SDQ mean total and subscale scores.

Although no interaction effect was observed for obesity status and sex on mean total scores, univariate F-tests indicated significant interactions on RSES and SDQ. Because there were significant interactions on the RSES and SDQ it was of interest to test which group was influencing the results, therefore F-tests were performed using individual mean squares error terms. F-tests confirmed this with obese girls reporting lower self-concept than their leaner peers on the RSES and the SDQ. These results concur with those reported by Martin et al (1988) and Drake (1985) who found obese girls to have lower self-esteem than their nonobese counterparts. Subscale scores for the PHCSCS indicated significant differences on the anxiety and popularity subscales with obese girls reporting statistically significantly lower mean scores than nonobese girls. For the SDQ, F-tests indicated significantly lower scores on general self and physical ability subscales for obese girls versus nonobese girls.

Overall, grouping obese boys and girls together can mask where the greater source of variance is being exhibited. To fully understand the relationship between obesity and self-esteem, gender differences need to be ascertained. These analyses have highlighted the dangers of grouping obese boys and girls data, as gender differences on subscale scores were influenced by obese girls. When gender was accounted for, the self-esteem of obese boys was comparable to that of their leaner peers. Low self-esteem both globally and in some subscales was confirmed for obese girls only. Analysis of the PHCSCS and the SDQ mean subscales scores indicated obese girls to be more anxious, less popular, and possess lower physical ability and general self-esteem. The result reported by Strauss et al (1985) such as rejection by peers is comparable to those noted for obese girls in this study who reported lower popularity self-esteem scores. It is difficult to ascertain the potential reasons for these data as no interviews were undertaken. However, one must be aware that the sociocultural pressures on females to be thin may influence self-esteem if one does not fit with this ideal (Kilbourne, 1994). The sociocultural pressures are further enhanced with the media images of the 'waif-like' popular culture that currently is in vogue. The increased attention and emphasis on 'ultra slimness' is even more likely to impact on obese individuals as they are further alienated from this 'popular ideal'. As Kilbourne (1994) suggested...

> "Adolescents are especially vulnerable, given the ominous peer pressure on young people. Also, normal physiological changes during adolescents result in increased body fat for women. If these normal changes are considered undesirable by culture (and by parents and peers) this can lead to chronic anxiety and concern about weight control in young women." (p. 397)

Thus, social rejection reported by obese girls may be a result of increased sociocultural pressure to be slim and to the stigma associated with obesity (Rothblum, 1995).

CHAPTER 5

Summary and Conclusions

Overall this study sought to determine whether measures based on the three theoretical models of self-esteem classified subjects in a similar manner. A secondary purpose was to determine the relationship between adolescent obesity and self-esteem by comparing the levels of self-esteem in this population with nonobese controls.

The internal and test-retest reliabilities of the PHCSCS and the SDQ were good and were comparable to values reported in the test manuals (Piers, 1984; Marsh, 1990a). Given the number of items in the RSES the reliability estimates were moderately strong but were slightly lower than those reported by Silber and Tippett (1965) and Fleming and Courtney (1984). The lower internal and test-retest reliability estimates obtained with the RSES suggested that it was the weaker of the three forms and that it was not as dependable indicator of self-esteem.

Loglinear analysis indicated that the three instruments were demonstrating conflicting patterns of relationships as no interaction significantly explained the best model to fit the data. The change in χ^2, obtained by removing the PHCSCS and SDQ interaction effect from the model, suggested that the RSES was contributing most to the poor fit of the model to the data. The lack of any interaction to adequately provide any pattern of relationships between instruments suggested the measures were not classifying individuals in a similar manner. However, correlations between all three measures suggested moderately high to high degrees of association. The lower correlations obtained between

the RSES and the PHCSCS and SDQ indicated a lower degree of association between the unidimensional measure and multidimensional and hierarchical measures. However, the loglinear analysis indicated that there was a high degree of association between the PHCSCS and the SDQ. In support of this, the correlation between the PHCSCS and the SDQ indicated that the multidimensional hierarchical model of self-esteem proposed by Shavelson et al (1976) to be plausible. Thus although there was no pattern of relationship between all three measures, correlations suggest multidimensional and hierarchical measures of self-esteem to be stronger indicators of self-esteem.

The multidimensional and hierarchical models of self-esteem had the advantage over the unidimensional model in that they provided evidence as to where low areas of self-esteem were manifest in obese subjects. The use of both multidimensional and hierarchical instruments provided a more in depth analysis to understanding self-esteem differences in obese and nonobese subjects. However, these results must be interpreted with caution given the results of the previous loglinear analysis. What these results provided though was a warning as to the dangers of grouping data for obese males and females, as evidence suggested obese girls to be the ones who reported lower self-esteem. It seems that self-esteem is not the same for obese males and females. Thus, in this particular sample, the condition of obesity had a differential impact on females than males. One cannot therefore take for granted that self-esteem is equivalent across gender for obese adolescents.

These analyses suggested that the self-esteem profile of obese girls warrants some concern, given the lower scores in psychosocial areas. Thus, the issue of peer rejection as

outlined by Stauss et al (1985), was evident for this group. It is difficult to ascertain from this study why obese girls reported lower self-esteem in these areas, given that paper and pencil self-report and no interviews were conducted. However, well known and established sociocultural pressures on females to be slim may adversely impact on the self-esteem of those who are not only not slim but, in fact are overweight (Kilbourne, 1994). The lower scores, specifically in the anxiety, popularity, and physical ability subscales for obese girls, are areas that require further examination. If low self-esteem is a mediating variable in the maintainance of obesity in girls, then appropriate interventions need to be targeted toward enhancing self-esteem in those areas that indicate lower self-esteem.

Suggested Future Research

The paucity in self-concept/esteem construct validation research noted by Byrne (1984), Hattie (1992), Wells and Marwell (1986) and Wylie (1974), is still apparent. The results from this study further emphasize the need for more construct validation research, given that the competing theoretical models of self-esteem failed to provide consistent patterns of relationships when classifying adolescent subjects.

Although the RSES is the most widely used self-esteem measure, there is surprisingly little construct validation research on this instrument. There is also a lack of published norms available for use with the RSES and therefore it is difficult to interpret the scores from this measure. If researchers are to continue to use this form as much as they have, then norm tables need to be established. Such tables could help provide guidelines for comparing self-esteem levels in children within and across samples.

The lower internal and test-retest reliabilities from the RSES suggested that there

was inconsistency in responding to the questions. The global nature of the questions may invite interpretation by the subjects. It is imperative to know how children are interpreting these global items therefore further research should ascertain the processes and experiences children utilize when making responses.

This research merely highlighted the differences in self-esteem between obese and nonobese subjects. Of more importance is to detail what the causes of low and high self-esteem are in order to further understand the relationship between adolescent obesity and self-esteem. The experiences, situations and significant others that impact on obese adolescents self-esteem need to be fully comprehended if meaningful interventions are to be developed and utilized. A more personal approach such as interview and observation in addition to sound psychometric measures might reveal the causes of how these impact on low/high self-esteem in this population.

These data revealed obese subjects to have lower self-esteem in certain subscale scores. Closer examination of the data indicated significantly lower self-esteem for obese girls only. Understanding what impacts on obese girls' self-esteem is of paramount importance as this may be a strong mediating variable in the onset and maintenance of obesity . Although obese boys presented a profile comparable to their leaner peers, future studies are required to replicate this finding. Additionally, research should focus on whether or not self-esteem structure is invariant across obese boys and girls given the results of this study.

Obesity status was derived from an objective measure of body composition. Actual obesity status may not be the key variable in the obesity and self-esteem

relationship, but rather self-perception of obesity level may be of primary importance. Thus attention should be directed towards perceptions of obesity and the associated levels of self-esteem.

Conclusion

Theoretical diversity as to how self-concept/esteem is ordered and measured has been a major retarding factor in self-concept/esteem research. No known research has attempted to assess whether self-esteem based on the three theoretical models classify individuals in a similar pattern. From a logical perspective, the functional worth of a measure must be demonstrated through its ability to provide similar patterns of results with other instruments purporting to assess the same construct. This study demonstrated that measures derived from the three theoretical models did not classify subjects in any consistent pattern. Therefore one cannot take for granted that the instruments used in this study were in fact measuring self-esteem. In line with Wylie's (1989) view it was concluded that the future use of these three measures warrants judicious interpretation of the results obtained and that they should not be used for diagnostic purposes.

The results from this study have highlighted the need to provide separate gender analyses for obese subjects, as obese girls were observed as having lower self-esteem globally and in certain subscale areas compared to obese boys. In order to fully understand the relationship between obesity and self-esteem it is important to document where differences in self-esteem occur within and across gender. The results from this study support this contention.

Unidimensionality has provided a platform on which self-concept/esteem research

has developed. However, recent self-concept/esteem theorists and researchers have advanced our understanding of self-perception through the development of more sophisticated scales that have embraced the Shavelson et al (1976) model. Given the growing number who are supporting the multidimensional model, it seems counterproductive to adopt undimensional measures to assess self-esteem. Unidimensional measures such as the RSES are quick and easy to administer, however, one should not compromise the quality and depth of self-perception assessment that can be gained from multidimensional and hierarchical measures, in favour of time and ease. The advantages of multidimensional and hierarchical measures are that they provide an in depth indication of where differences in self-esteem occur and allow for relationships between subscales and the construct they reflect to be examined within and across samples. Marsh and Shavelson (1985) stated that self-concept/esteem cannot be fully understood if its multidimensionality is ignored. The results of this study support this contention.

References.

Allon, N. (1979). Self-perceptions of the stigma of overweight in relation to losing weight patterns. The American Journal of Clinical Nutrition, 32, 470-480.

Blascovich, J. & Tomaka, J. (1991). Measures in self-esteem. In J. P. Robinson, P. R. Shaver & L. S. Wrightsman (Eds). Measures of Personality and Social Psychological Attitudes. Vol 1 Academic Press.

Bracken, B. A. (1988). Ten psychometric reasons why similar tests produce dissimilar results. Journal of School Psychology, 26, 155-166

Bracken, B. A. & Howell, (1991). Multidimensional self-concept validation: A three instrument investigation. Journal of Psycoeducational Assessment, 9, 319-328

Burton, B. T., Foster, W. R., Hirsch, J. & Van Itallie T. B. (1985). Health implications of obesity: an NIH Consensus Development Conference. International Journal of Obesity, 9, 155-169.

Byrne, B. M. (1983). Investigating measures of self-concept. Measurement and Evaluation in Guidance, 16, (3), 115-126.

Byrne, B. M. (1989). Multigroup comparisons and the assumption of equivalent construct validity across groups: Methodological and substantive issues. Multivariate Behavioral Research, 24 (4), 503-525.

Campbell, D. T. & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. Psychological Bulletin, 56, (2), 81-105.

Canada Fitness Survey. (1985). Canadian Youth Fitness. Ottawa: Fitness Canada, Fitness and Amateur Sport.

Canadian Standardized Test of Fitness (CSTF). (1986). Operations Manual. 3rd ed. Ottawa: Minister of State, Fitness and Amateur Sport.

Cook, T. D. (1979). Quasi-experimentation: Design & analysis issues for field settings. Chicago: Rand McNally.

Cook, T. D. & Campbell, D. T. (1979). Design and Analysis Issues for Field Settings. Chicago: Rand McNally.

Cooley, C. H. (1902) Human Nature and Social Order. New Brunswick (U.S.A.) and London (U.K.):Transaction Books. (1983)

Coopersmith, S. (1959). A method for determining types of self-esteem. Journal of Abnormal and Social Psychology, 59, 87-94.

Coopersmith, S. (1967). The antecedents of self-esteem. San Francisco: Freeman.

Corbin, C. B., Pangrazi, R. P., Corbin, W. R., Peterson, G., Pangrazi, D. & Meza, A. Z. (1995). Self-esteem profiles: A comparison of boys above below the national criteria for body fatness. Medicine and Science in Sports and Exercise, 27, (5) S115 (Supplement)

Cronbach, L. J. & Meehl, P. E. (1955). Construct validity in psychological tests. Psychological Bulletin, 52, (4), 281-302.

Delugach, R. R., Bracken, B.A, Bracken, M. J. & Schicke, M. C. (1992). Self-concept: Multidimensional construct interpretation. Psychology in Schools, 29 213-223.

Demo, D. H. (1985). The measurement of self-esteem: Refining our methods. Journal of Personality and Social Psychology, 48 (6), 1490-1502.

Demo, D. H. & Savin-William, R. C. (1983). Early adolescent self-esteem as a function of social class. American Journal of Sociology, 88, 763-774.

Drake, M, A. (1988). Self-esteem of adolescents enrolled in a weight reduction program. Journal of the American Dietetic Association, 88 (12), 1581-1582.

Flemming, J. S. & Courtney, B. E. (1984). The dimensionality of self-esteem:Hierarchical facet model for revised measurement scales. Journal of Personality and Social Psychology, 46 (2), 404-421.

Fox, K. R. (1988). The self-esteem complex. British Journal of Physical Eduction, 19, (6), 247-252.

Fox, K. R. (1992). Physical education and the development of self-esteem in children. In N. Armstrong (Ed) New Directions in Physical Education Vol 2. Towards a National Curriculum.

Fox, K. R. & Corbin, C. B. (1989). The physical self-esteem profile: Development and preliminary validation. Journal of Sport and Exercise Psychology, 11, 408-430.

Garn, S. M., Leonard, W. R. & Hawthorne, V. M. (1986). Three limitations of the body mass index. American Journal of Clinical Nutrition, 44, 996-997.

Gortmaker, S. L., Deitz, W. H., Sobol, A. M. & Wehler, C. A. (1987). Increasing pediatric obesity in the United States. American Journal of deceases of Children, 141, 535-540.

Gruber, J. J. (1986). Physical activity and self-esteem development in children: A meta analysis. The American Academy of Physical Education. The Academy Papers, The Effects of Physical Activity on Children, 30-48.

Hamilton, D. L. (1971). A comparative study of five methods of assessing self-esteem dominance and dogmatism. Educational and Psychological Measurement, 31, 441-452.

Hattie, J. (1992). Self-Concept. Hillsdale, New Jersey: Lawrence Erblaum Associates, Publishers.

Hattie, J., & Marsh, H. W. (in press). Future directions in self-concept research. In B. A. Bracken (ed.), Handbook of Self-Concept. New York, NY: Wiley.

Israel, A. C. & Shapiro, L. S. (1985). Behaviour problems of obese children enrolling in a weight reduction program. Journal of Pediatric Psychology, 10 (4), 449-460.

James, W. (1890). The Principles of Psychology (Vol 1). Dover Authorized Edition, 1950, New York: Dover Publications

Kaplan, K. M. & Wadden, T. A. (1986). Childhood obesity and self-esteem. Journal of Pediatrics, 109 (2), 367-370.

Kennedy, J. J. (1983). Analyzing Qualitative Data. Introductory Loglinear Analysis for Behavioral Research. New York, NY: Praeger.

Kilbourne, J. (1994). Still killing us softly: Advertising and the obsession with thinness. In P. Fallon, M. A. Katzman & S. C. Wooley (Ed.) Feminist Perspectives on Eating Disorders. New York, NY: Guilford.

Kimm, S. Y. S. Sweeney, C. G. & Janosky, J. E. (1991). Self-construct measures and childhood obesity: A descriptive analysis. Developmental and Behavioral Pediatrics, 12 (1), 19-24.

Klesges, R. C., Haddock, C. K., Stien, R. J., Klesges, L. M., Eck, L. H. & Hanson, C. L. (1992). Relationship between psychosocial functioning and body fat in preschool children: A longitudinal investigation. Journal of Consulting and Clinical Psychology, 60, (5), 793-796.

Lerner, R. M. & Gellert, E. (1969). Body build identification, preference and aversion in children. Developmental Psychology, 1 (5), 456-462.

Lohman, T. G. (1992). Advances in Body Composition Assessment. Champaign, Illinois: Human Kinetics Publishers.

Maguire, T. O., Hattie, J., & Haig, B. (1994). Construct validity and achievement assessment. The Alberta Journal of Educational Research. Vol XL (2), 109-126.

Manus, H. E., & Killeen, M. R. (1995). Maintenance of self-esteem by obese children. Journal of Child and Adolescent Psychiatric Nursing, 8, (1), 17-28

Marsh, H. W. (1987). The hierarchical structure of self-concept and the application of hierarchical confirmatory factor analysis. Journal of Educational Measurement, 24 (1), 17-39.

Marsh, H. W. (1990 a). Self Description Questionnaire(SDQ) II Manual. Publication Unit, Faculty of Education, University of Western Sidney, Macarthur.

Marsh, H. W. (1990 b). Confirmatory factor analysis of multitrait-multimethod data. The construct validation of multidimensional self-concept measures. Journal of Personality, 58, (4), 661-692.

Marsh, H. W., & Shavelson, R. (1985). Self-concept: Its multifaceted, hierarchal structure. Educational Psychology, 20 (3), 107-123.

Marsh, H. W. & Smith, I. D. (1982). Multitrait-multimethod analyses of two self-concept inventories. Journal of Educational Psychology, 74, (3), 430-440.

Marshall, J. D., Hazlett, C. B., Spady, D. W. & Quinney, H. A. (1990). Comparison of convenient obesity indicators. American Journal of Clinical Nutrition, 51, 22-28.

Marshall, J. D., Hazlett, C. B., Spady, D. W., Conger, P. R. & Quinney, H. A. (1991). Validity of convenient indicators of obesity. Human Biology, 63, (2), 137-153.

Martin, S. Housley, K. McCoy, H. Greenhouse, P. Stigger, F. Kenney, M. A. Shoffner, S. Fu, V. Korsland, M. Ercanli-Huffman, E. G. Carter, E. Chopin, L. Hegsted, M. Clark, A. J. Disney, G. Moak, S. Wakefield, T. & Stallings, S. (1988). Self-esteem of adolescent girls as related to weight. Perceptual and Motor Skills, 67, 879-884.

McLaren, D. S. (1987). Letters to the editor. Three limitations of the body mass index. American Journal of Clinical Nutrition, 46, 121-123.

Mead, G. H. (1967)(1934). Mind, Self and Society: From the Standpoint of a Social Behaviourist. Chicago and London: The University of Chicago.

Mendolson, B. K. & White, D. R. (1982). Relation between body self-esteem and self-esteem of obese and normal children. Perceptual and Motor Skills, 54, 899-905.

Messick, S. (1989). Validity. In R. L. Linn (Ed), Educational measurement. New York, N.Y : American Council on Education.

O'Brien, R. Smith, S. A. Bush, P. J. & Peleg, E. (1990). Obesity, self-esteem and health locus of control in black youths during transition to adolescence. American Journal of Health Promotion, 5 (2), 133-138.

Pedhazur, E. J. & Schmelkin, L. J. (1991). Reliability. Measurement, Design, and Analysis: An Integrated Approach. Hillsdale, N J: Lawrence Erlbaum Associates.

Piers, E. V. (1984). Piers-Harris Children's Self-Concept Scale: Revised Manual. Los Angeles, CA: Western Psychological Services.

Rogers, W. T. (1994). The treatment of measurement issues in the revised program evaluation standards. Paper presented at the Annual Meetings of the American Educational Research Association.

Rosenberg, M. (1979). Conceiving the Self. New York: Basic Books

Rothblum, E. D. (1994). "I'll die for the revolution but don't ask me not to diet:Feminism and the continuing stigmatization of obesity. In P. Fallon, M. A. Katzman & S. C. Wooley (Ed.) Feminist Perspectives on Eating Disorders. New York, NY: Guilford.

Sallade, J. (1973). A comparison of the psychological adjustment of obese vs non-obese children. Journal of Psychosomatic Research, 17, 89-96.

Savin-Williams, R. C. & Jaquish, G. A. (1981). The assessment of adolescent self-esteem: A comparison of methods. Journal of Personality, 43, (3), 324-336.

Shavelson, R. J. & Bolus, R. (1982). Self-concept: The interplay of theory and methods. Journal of Educational Psychology, 74, (1), 3-17.

Shavelson, R. J. Hubner, J. J. & Stanton, G. C. (1976). Self-concept: Validation of construct and interpretations. Review of Educational Research, 46 (3), 407-441.

Silber, E. & Tippett, J. S. (1965). Self-esteem: Clinical assessment and measurement validation. Psychological Reports, 16, 1017-1071.

Smith, M. L. & Glass, G. V. (1987). Research and Evaluation in Education and the Social Sciences. Boston: Allyn and Bacon.

Strauss, C. C. Smith, K. Frame, C. Forehand, R. (1985). Personal and interpersonal characteristics associated with obesity. Journal of Pediatric Psychology, 10 (3), 337-343.

Traub, R. E. (1994). Measurement Methods for the Social Sciences. Reliability for the Social Sciences: Theory and Apllication (Vol 3). Thousand Oaks: Sage.

Van Tuinen, M. & Ramanaiah, N. V. (1979). A multimethod analysis of selected self-esteem measures. Journal of Research in Personality, 13, 16-24.

Wadden, T. A. & Stunkard, A. J. (1985). Social and psychological consequences of obesity. Annals of Internal Medicine, 103 (6), 1062-1067.

Wadden, T. A. Foster, , G. D. Brownell, K. D. & Finley, E. (1984). Self-concept in obese and in normal weight children. Journal of Consulting and Clinical Psychology, 52 (6), 1104-1105.

Wells, L. E. & Marwell, G. (1976). Self-esteem: Its Conceptualization and Measurement. London: Sage Publications.

Williams, M.H. (1986). Weight control through exercise and diet for children and young athletes. The American Academy of Physical Education. The Academy Papers, The Effects of Physical Activity on Children. 88-113.

Winnie, P. H. & Marx R. W. (1977). A multitrait-multimethod study of three self-concept inventories. Child development, 48, 893-901.

Winnie, P. H. & Marx R. W. & Taylor T. D. (1977). A mulitrait-multimethod study of three self-concept inventories. Child Development, 48, 893-901.

Wylie, R. C. (1974). The Self Concept: Revised Edition (Vol 1). Lincoln: University of Nebraska Press.

Wylie, R. C. (1989). Measures of Self Concept. Lincoln: University of Nebraska Press.

APPENDIX A

Consent Form

CONSENT FORM

Dear Parent(s)/Guardian(s):
Researchers at the Faculty of Physical Education and Sports Studies at the University of Alberta are undertaking studies of self-esteem in relation to body composition.

As Principal Investigators of this project we are planning to conduct a measurement study in _____during the winter/ spring of 1995, the primary purpose of which is to determine the relationship between self-esteem and body composition. Your child has been one of 500 children who has randomly been selected to participate in the study.

Children who receive consent from their parent(s)/guardian(s) will be asked to complete three paper and pencil self-esteem inventories and will have body composition measured. The self-esteem questionnaires are standard forms that are widely used with children of this age range and will take approximately 40 minutes to complete. Body composition measures will include height, weight, 5 skinfolds thickness measurements and a subjective visual rating of the child's body fat level. All measures will be taken during the physical education classes, so there will be minimal disruption to your child's daily school routine.

All of the measures to be taken are standard measures; there will be virtually no risk to the child. All records will be kept strictly confidential. The child will learn about self-esteem and body composition. The knowledge gained will allow the investigators to understand better the relationship between body composition and self-esteem.

If you consent to having your child participate in this study, please complete the enclosed form. If you have any questions regarding the study, please feel free to call us on the numbers listed on the consent form. Thank you for your consideration.

Sincerely,

Dru Marshall, Ph.D. Brian Nielsen, Ph.D. Richard Fletcher BA(Hons).

TITLE:Body Composition and Self-Esteem: A Validation of Selected Self-Esteem Measures

I hereby certify that_____(child's name), for which I am the parent/guardian, is allowed to participate in the research study directed by Drs. Dru Marshall, Brian Nielsen and Mr R Fletcher, that will take place within_____(schools name) during the physical education class. In doing so, I understand fully all of the statements below:

1. The study will involve 500 students where body composition and self-esteem data will be collected. Body composition measures will include height, weight, a subjective visual rating of the students body fat level and 5 skinfolds thickness measurements(Front, back of the arm, below the shoulder blade, above the hip, on the inside of the calf). Self-esteem will be measured using three paper and pencil questionnaires and will take approximately 40 minutes to complete. I understand that there will be virtually no danger or risk in the collection of these measures.

2. The disruption to the child's normal school routine will be minimal, as all measures will occur during scheduled physical education class time.

3. I agree that my child will voluntarily participate in the study as it is described. I understand that my child has the right to withdraw from the study at any time with no penalty. I understand that there is no financial renumeration for participating in this study.

4. My child's identity will not be disclosed during the time of his/her participation in the future, or in any published results.

5. I understand that if I have any questions related to any part of my child's participation in this project, my questions will be fully answered to my total satisfaction.

6. I hereby make available to Drs. Marshall, Nielsen and Mr Fletcher at the University of Alberta all results obtained as a consequence of my child's participation in this project, whether these results are in individual or group form.

7. I further certify that all procedures in which my child will be involved have been fully explained to me, and will be fully explained to my child. Any rights have also been explained to me. I hereby declare that I am totally satisfied with these explanations.

8. I understand that this study will further research aimed at self-esteem and body composition.

The persons who may be contacted about the research are:

Dr. Dru Marshall: 492-1035 Dr. Brian Nielsen:492-3839 Mr. Richard Fletcher: 434-0489

_____ (Date) _____(Signature of Investigator or Designee)

_____(Name) _____(Signature of Parent/Guardian)

_____(Name of Witness) _____(Signature of Witness)

APPENDIX B

Rosenberg Self-Esteem Scale (Rosenberg, 1979)

The Piers-Harris Children's Self-Concept Scale (Piers, 1984)

The Self-Description Questionnaire II (Marsh 1990)

ROSENBERG SELF-ESTEEM SCALE
Morris Rosenberg 1979

Respondents are asked to strongly agree,agree, disagree, or strongly disagree with the following items.

1.	On the whole I am satisfied with myself	SA	A	D	SD
2.	At times I think I am no good at all	SA	A	D	SD
3.	I feel that I have a number of good qualities	SA	A	D	SD
4.	I am able to do things as well as most other people	SA	A	D	SD
5.	I feel I do not have much to be proud of	SA	A	D	SD
6.	I certainly feel useless at times	SA	A	D	SD
7.	I feel that I'm a person of worth, at least on an equal plane with others	SA	A	D	SD
8.	I wish I could have more respect for myself	SA	A	D	SD
9.	All in all, I am inclined to feel that I am a failure	SA	A	D	SD
10.	I take a positive attitude toward myself	SA	A	D	SD

"THE WAY I FEEL ABOUT MYSELF'

The Piers-Harris Children's Self-Concept Scale

Published by:
Western Psychological Services
Publishers and Distributors
12031 Wiltshire Boulevard
Los Angeles, California 90025-1251
Copyright © 1969 Ellen V Piers and Dale B Harris
Not to be reproduced in whole or part without written permission Western Psychological Services. All right reserved

Due to copyright © laws pertaining to the Piers-Harris Children's Self-Concept Scale, no items may be reproduced without written consent of the publishers, Western Psychological Services.

THE SELF-DESCRIPTION QUESTIONNAIRE-II
Herbert W Marsh 1990

Published by the:
Publication Unit
Faculty of Education
University of Western Sydney
Macarthur
PO Box 555
Campbelltown
N.S.W. 2560
Australia.

Permission to use this form can be obtained by contacting Professor H.W. Marsh at the University of Western Sydney, Macarthur, PO Box 555, Campbelltown, N.S.W. 2560, Australia.